THE PROCEEDINGS OF
THE PHYSICAL SOCIETY
Section B

Vol. 65, Part 11 1 November 1952 No. 395 B

. B
5
2

X-Ray Analysis of Proteins

By W. L. BRAGG
Cavendish Laboratory, Cambridge

36th Guthrie Lecture, delivered 12th March 1952; *MS. received 9th June* 1952

ABSTRACT. The attempt to discover the atomic arrangement in the protein molecule seems very ambitious. Ever since Bernal first showed that crystals of protein gave x-ray diffraction pictures, it has been clear that a protein molecule of a given type is a structure with a definite individual form; the x-ray diffraction spots are very sharp and reproducible and extend to regions corresponding to interatomic distances. The molecules are, however, of great complexity. It has been a triumph of x-ray analysis to pass from simple substances like rock salt to such molecules as strychnine or penicillin with about one hundred atoms. We are now trying to analyse a molecule such as haemoglobin, which contains ten thousand atoms. The reward, if an analysis were completed, would be great, because the determination of any one protein would undoubtedly cast a flood of light on the character of these bodies, which Nature has selected as the basis of living matter.

Work on protein is going on in a number of centres. I wish to speak here in particular about the work on haemoglobin which has been going on in Cambridge for the last twelve years under the direction of Dr. Perutz. More x-ray data for haemoglobin are available than for any other protein. Though still a long way short of the goal, considerable advances have been made recently, so great as to encourage us to hope that the problem is soluble and even that a solution lies not far ahead.

---

§ 1.

The globular proteins are very large molecules which form an essential part of living matter. It would be of intense interest to find out their structure. In this lecture I propose to describe one line of attack on this problem, the analysis of the x-ray diffraction pictures given by these proteins in a crystalline form. The aim of x-ray analysis is to provide a picture of the arrangement of the atoms in matter of different kinds, and a knowledge of this arrangement greatly strengthens our understanding of the structure and properties of the bodies. In the case of the silicate minerals, for example, the determination of the structures brought order into a very confused pattern, and I think it is fair to say that it made the relationship between the minerals clear for the first time. Inorganic stereochemistry had hardly even been the subject of speculation until x-ray analysis was successful in supplying information about the relative positions of the atoms in inorganic compounds. It then became clear that the stereochemistry of these bodies is governed by very elegant and simple laws leading to a minimal energy of the field between charged ions, as was first shown by Pauling. Even in the organic field, where stereochemistry has been studied with such striking success by the chemists, x-ray analysis is clearing up doubtful cases and adding precision to the whole picture. It is to be hoped therefore that a much better insight into the nature and function of the molecules in living matter will be obtained if x-ray analysis is able to reveal their atomic arrangement. This is the prize which leads the investigator on to attempt the analysis of these extremely complex bodies.

I am interested in the application to this problem of physical methods based on optical principles, and this is the side in which I hope to interest you. A good deal of work has been done by the biochemists on the structure of these substances. If I refer to it, it is merely because it supplies essential information which helps us to survey our problem. I am no chemist, and my references to the chemical side are only made because it helps to provide a starting point for the physical investigations.

§2.

The size of the protein molecule is so great that it is impossible to attack its structure by the ordinary methods of x-ray analysis. The normal procedure is to make a guess at the arrangement of the atoms in the crystal and check whether this postulated arrangement would diffract the x-rays in the manner actually observed. The guesses are inspired guesses based on previous experience of the relative arrangements of atoms, on what the chemists can tell us about the molecule in the case of an organic compound, and on any other scrap of information which narrows the range of likely structures, but it remains a guess. X-ray analysis is essentially a cut and try method. It is remarkable what very complex structures can be analysed with these limitations. Perhaps the most striking instance in recent years was the successful x-ray analysis of penicillin by Mrs. Hodgkin at Oxford and Dr. C. W. Bunn at I.C.I. A number of pieces of this molecule had been definitely established by the chemists, and guesses had to be made as to how they joined together. Even with this help the analysis took several years and represented a remarkable achievement. Now the molecular weight of penicillin is rather over 300. To take some cases of typical protein molecules, the molecular weight of insulin is 12000, that of myoglobin 17000, that of pepsin 37000 and that of haemoglobin 66700. The cut and try method in molecules containing so many atoms is out of the question; we must proceed by a more direct method in which guess-work plays as little part as possible.

The chemists tell us that these proteins are built of the residues of amino acids. The nature of these bodies is illustrated in fig. 1. A typical amino acid is built by starting with an asymmetric carbon atom (the α carbon) and attaching to it tetrahedrally a —COOH group, a —NH_2 group, a hydrogen atom, and a fourth group which gives the acid its specific character. Examples of these groups are shown in the upper part of fig. 1. They vary from being simply a hydrogen atom to being quite a large tag, like tryptophan. On the average the tag or labelling group contains about four atoms, so it is not very complex. These amino acids can polymerize into long chains, the polypeptide chains. The

—COOH and the —NH_2 groups are like male and female joints. A reaction takes place between the —COOH of one acid and the —NH_2 of another in which water is eliminated and a firm chemical bond is made. A polypeptide chain is shown diagrammatically in the lower part of fig. 1. These chains are the building stuff of the proteins, and the particular nature of the protein appears to be determined by the order in which the different kinds of amino acid are linked together. There are rather over twenty different kinds of amino acid which occur in all natural bodies. One may think of them as being like the letters of the alphabet joined together in a sequence which gives significance and character to a given type of protein. It is interesting to note that all these amino acids appear to be left-handed in the sense of the chemical convention. Nature would work as well if they were all right-handed, but it would appear that some chance

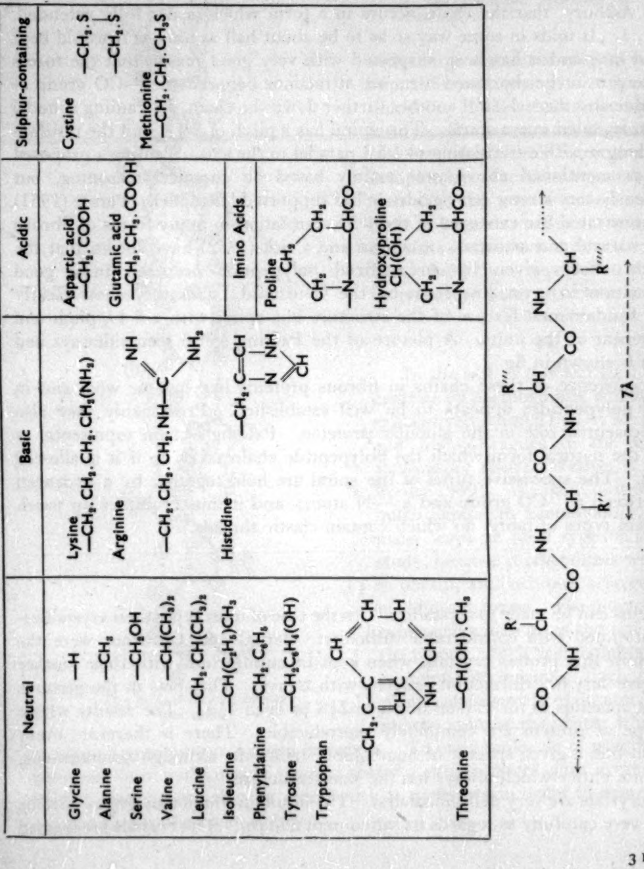

Neutral
Glycine —H
Alanine —CH_3
Serine —CH_2OH
Valine —$CH(CH_3)_2$
Leucine —$CH_2.CH(CH_3)_2$
Isoleucine —$CH(C_2H_5)CH_3$
Phenylalanine —$CH_2.C_6H_5$
Tyrosine —$CH_2C_6H_4(OH)$
Tryptophan
Threonine —$CH(OH).CH_3$
Basic
Lysine —$CH_2.CH_2.CH_2.CH_2(NH_2)$
Arginine —$CH_2.CH_2.CH_2.NH$—C
Histidine
Acidic
Aspartic acid —$CH_2.COOH$
Glutamic acid —$CH_2.CH_2.COOH$
Imino Acids
Proline
Hydroxyproline
Sulphur-containing
Cystine —$CH_2.S$
Methionine —$CH_2.CH_2.CH_3S$
R′
R″
R‴
R⁗
7Å

Fig. 1. An amino-acid has a central carbon atom to which are attached a —COOH group, an —NH_2 group, a hydrogen atom, and one of the character-giving groups shown in the upper part of the figure. The acids join into long polypeptide chains by the —COOH group of one reacting with the —NH_2 group of the next with the elimination of water, as shown in the lower part of the figure.

event has started all living matter on the left-handed convention. We are irrevocably left-handed. The taste and smell of right- and left-handed organic bodies is quite different to us, although they are completely identical in structure, and foods or poisons are similarly distinct.

A protein molecule contains a large number of these amino acid residues. For instance there are 106 of them in insulin, 146 in myoglobin, and 540 in haemoglobin. Exhaustive chemical analysis has in certain cases yielded information about the successive order of the residues along the chains, as for instance has been done for insulin by Sanger.

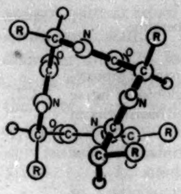
Quite recently Pauling in America has proposed a model for the actual form of the polypeptide chain based on chemical reasoning. The bond lengths and bond angles in simple amino acids have been well established, and Pauling bases his model on these data. It has long been evident, particularly from the work of Astbury, that the chain occurs in a form which is not fully extended, as in fig. 1. It folds in some way so as to be about half as long as it would be if stretched out, and it has been suspected with very good reason that the forces which keep it in the shortened form are attractions between the —CO group of one residue and the —NH of another further down the chain. In Pauling's model the chain is coiled into a spiral. This spiral has a pitch of 5·4 Å, and the residues repeat along it with a translation of 1·5 Å parallel to the axis. Pauling's proposed model, as mentioned above, was mainly based on chemical reasoning, but subsequently very strong x-ray evidence has supported his model. Perutz (1951) has demonstrated the existence of the 1·5 Å translation in many forms of fibrous protein, natural and artificial. Cochran and Crick (1952) have shown that the diffraction effects given by an artificial polypeptide prepared in a good approximation to crystalline form by the Courtauld Laboratory show clearly that the fundamental feature of the structure is a spiral with a 5·4 Å pitch and a 1·5 Å repeat of the units. A picture of the Pauling spiral seen sideways and endways is shown in fig. 2.

The existence of these chains in fibrous proteins like hair or wool and in artificial polypeptides appears to be well established. Presumably they also play an essential role in the globular proteins. Pauling's chain represents, as it were, the natural form which the polypeptide chain takes up if it is allowed to do so. The successive turns of the spiral are held together by a hydrogen bond between a —CO group and a —N atom; and it thus crumples up much like certain types of fabric do which contain elastic threads.

§ 3.

Proteins can be made to crystallize. In the case of most types their crystallization is attended with considerable difficulty. Bernal and Crowfoot were the first to show that protein crystals, when kept in equilibrium with their mother liquor, give very fine diffraction patterns with x-rays. The spots in the pictures represent spacings in the crystal down to 2½ Å or even 1½ Å. The results with a given type of protein are completely reproducible. There is therefore every indication that a given species of homoglobin molecule, although so enormous, is a definite entity which always has the same structure.

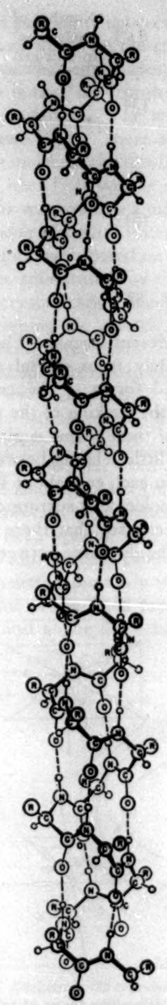
The crystals are very delicate affairs. The liquid in which they grow has to be adjusted very carefully as regards its saltcontent and pH. The crystals are formed



Plan of the 3·7-residue helix.

Fig. 2. Model of the spiral polypeptide chain proposed by Pauling, Corey and Branson (1951).

Reproduced by permission from the 'Proceedings of the National Academy of Sciences', Washington, D.C.



The helix with 3·7 residues per turn.

with a large amount of water of crystallization, often occupying more than half the unit cell. This water can be withdrawn in successive stages and, if great precautions are taken, it can be removed almost completely. One has the general impression that the crystal consists of molecules of large size with a complex pattern of points of attachment over their surface which hold them together in the crystal by relatively weak attractive forces.

§ 4.

In the rest of this lecture I will concentrate on haemoglobin, which has been exhaustively studied by Dr. M. F. Perutz and his colleagues in the Cavendish Laboratory (Boyes-Watson *et al.* 1947, Perutz 1949). We have, as it were, sunk much capital in haemoglobin. It is in many ways an ideal type of protein to study, because it crystallizes well, is easy to obtain, and exhibits a large number of different forms. Haemoglobin is the substance which gives the red colour to the blood. Each haemoglobin molecule contains four haems. Each haem is a planar arrangement of organic rings with an iron atom at the centre which is able to take up or release an oxygen molecule. To use a molecule of weight 66 700 to transport four molecules of oxygen seems a rather extravagant arrangement of Nature, but presumably its complexity is necessary in order to ensure that the oxygen gets to the right address.

Substance	Space group	*a*	*b*	*c*	β (°)	Number of molecules in unit cell	Vol. per molecule (Å^3)
a. Met-, oxy-, carboxy-haemoglobin of horse	C2	102	56	49	134	2	103000
b. Reduced haemoglobin of horse	$C3_12$	47·4	—	308		6	100000
c. Methaemoglobin of foetal sheep. I	$B22_12$	78	99	54		4	104000
d. Methaemoglobin of foetal sheep. II	$P2_1$	48	71	58	101	2	97000
e. Met-, oxy- carboxy-haemoglobin of man	$P4_12_1$	47·4	—	174		4	98000
f. Reduced haemoglobin of man	$P2_1$	59	70	47·5	82·5	2	97000

Haemoglobin has been crystallized in many different forms. The amount of water associated with the molecule varies widely from crystal to crystal in what we call the 'wet' forms, but when dried forms of the crystal can be obtained without distorting the whole structure the packing of the molecules into these forms is of great interest. A list of six of these forms is given in the table. Crystallographically there seems to be very little relation between them, but it will be noted that the volume per molecule is in each case about 100 000 Å^3. If now we draw pictures of the way in which one molecule is surrounded by its neighbours in these different crystalline forms we can see that there are close relationships. This has been done in fig. 3, where the distances between nearest neighbours and the directions in which they lie are indicated. Although the crystalline forms appear so different they are actually distorted versions of cubic or hexagonal close packing. Each protein molecule touches twelve neighbours,

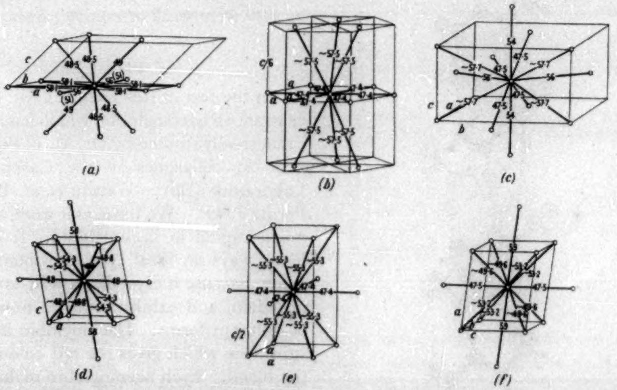


Fig. 3. Distances between neighbouring molecules in various crystalline forms of haemoglobin. They suggest that the 'dry' molecule is roughly egg-shaped with a diameter of 47 Å and a length of 57 Å.

as in close packing of spheres. These neighbours are not quite at equal distances, however, as would be the case if the protein molecule were round. It appears to be somewhat egg-shaped, and about 47 Å in diameter and 57 Å long. This of course is the 'dry' protein molecule from which as much water as possible has been removed. The water associated with the 'wet' forms would appear to be of two kinds. There appears to be a layer of bound water, on the average about one molecule thick, which is in some way closely associated with the protein molecule itself. It will be realized that these are very tentative ideas: we are feeling our way towards a picture of the molecule. Outside this protein molecule with its layer of bound water there is the water of crystallization in the unit cell. This distinction is based on an ingenious argument due to Perutz. When a salt solution is substituted for the water surrounding the crystal the salt enters the water of crystallization inside the crystal, and one can measure how much goes in by weighing the crystal. The substitution of salt solution for water causes an interesting X-ray diffraction change. The low orders of X-ray diffraction near the central spot are due to the protein molecule as a whole being higher in density that the water in which it swims. By increasing the salt concentration these low orders can be made to disappear. We have in fact matched the X-ray density of the protein by that of the solution, just as optical densities are matched by changing the liquid when measuring the refractive index of a small crystal. Since the amount of salt entering the unit cell is known, and since we know that the salt solution must match the protein, we can say through what volume this salt has been spread, and so deduce the volume associated with the protein molecule into which it cannot enter. This turns out to be 116 000 $Å^3$, the volume of the hydrated molecule with a layer of water all round it, and not 83 000 $Å^3$, the volume of the dried protein molecule. The bound water seems to keep out the salt and so be something especially associated with the protein molecule (fig. 4).

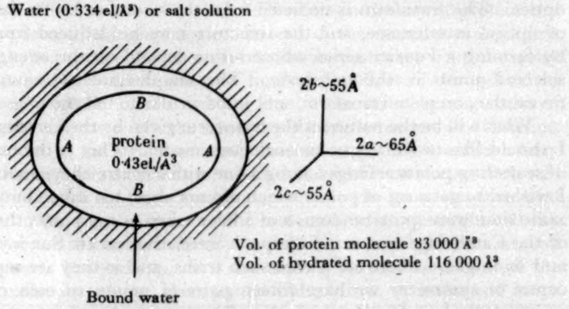

Water (0·334 el/Å³) or salt solution
B
A
protein
0·43el./Å³
A
B
2b~55Å
2a~65Å
2c~55Å
Vol. of protein molecule 83 000 Å³
Vol. of hydrated molecule 116 000 Å³
Bound water

Fig. 4. Diagrammatic representation of the haemoglobin molecule surrounded by a layer of bound water.

Figures 4 to 9 are reproduced by permission from the 'Proceedings of the Royal Society' (Bragg and Perutz 1952).

§ 5.

Let us now consider what type of diffraction effects such a molecule will give. At the degree of resolution which we are considering the water or salt solution may be thought of as completely uniform and not giving rise to diffraction. Its spacings are very small, of the order of 2 Å or 3 Å, and we can concentrate here on X-ray diffraction effects due to the larger spacings which will outline the general structure of the molecule. We therefore have a protein molecule in a uniform sea of water or salt solution. It is heavier than the water, and it contains detail in the way of greater and lesser concentrations of scattering power due to its structure.

In the case of an ordinary optical line grating it will be remembered that the positions of the different orders of spectra are determined by the spacing of the lines, whereas the relative strengths of the orders are determined by the nature of an individual line. They are strong if the typical line scatters strongly in that direction, and weak if the diffraction by the line is weak. I use this analogy to introduce an idea which is one of the basic ones in X-ray diffraction. To any one kind of scattering unit or unit of pattern in the crystal can be associated a 'transform'. We may think of this as a sort of three-dimensional Fraunhofer diffraction pattern which records how strongly that unit diffracts in any particular direction when we observe X-ray diffraction effects by the crystal. The positions of the spots depend of course on the crystal spacings, but the strength of a spot depends on whether the transform has a high or low value at that place. One is familiar with this transform idea in the case of Fraunhofer diffraction; for instance, the transform of a slit is a central positive peak with maxima and minima on either side which alternate in sign and rapidly diminish. The transform of a circular aperture is a circular positive peak with a negative ring around it and then a still weaker positive ring, and so forth. In X-ray analysis, by measuring the strength of the diffracted spots we are, as it were, obtaining samples of the transform at these places. If only we knew the amplitude and phase throughout the transform we could immediately deduce the structure that produced it. Structure and transform are related in the reciprocal way so familiar in physical optics. The transform is deduced from the structure by the familiar principles of optical interference, and the structure may be deduced from the transform by forming a Fourier series whose terms depend on the strength and phase at selected points in the transform. The samples are not enough, however, to reveal the complete transform, and in particular to tell the phase at each point.

What will be the nature of the transform given by the haemoglobin molecule? I should like to show you an analogous model. This is the transform of a set of scattering points arranged along a line with a centre of symmetry in the middle. I wished to get a set of points which did not show too much bunching and at the same time were quite random, and I have taken as my points the times of arrival of the Cambridge trains at Liverpool Street Station on Sunday between 8 a.m. and midnight. There are fifteen such trains, and as they are repeated about the centre of symmetry we have fifteen pairs of points, to each of which I have assigned a value of unity.

The transform of this group, actually in two forms obtained by putting the centre of symmetry at A or at B, is shown in fig. 5, and I would like to draw your attention to the following points of interest. It has of course a large central positive maximum corresponding to the rays scattered through a very small angle. It then has a minimum, followed by a maximum. The positions and heights of these are very closely those which would be given by a continuous strip of scattering matter of width 2AB. At this low resolution the exact positions

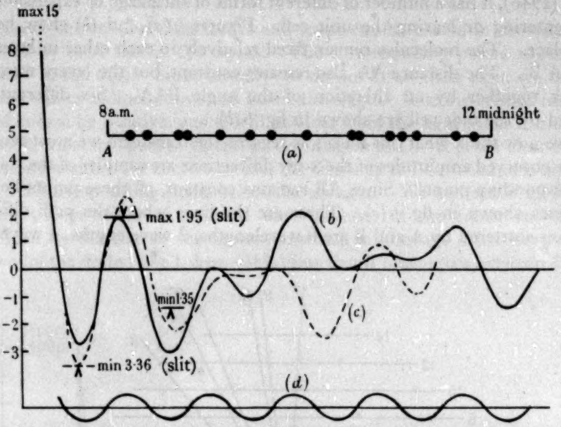


Fig. 5. The transform of a series of scattering points within a limited space.

(*a*) Set of 15 random points between A and B.

(*b*) Transform of points in pairs, each pair of weight unity, along a line BAB with centre of symmetry at A.

(*c*) Transform for ABA with centre of symmetry at B.

(*d*) Positions (not amplitudes) of Fraunhofer maxima and minima for slit of width 2AB. The amplitudes are indicated for comparison with curves (*b*) and (*c*).

of the points are unimportant; we have merely got a region of high density smeared out over the strip. This corresponds to the feature in the protein that the lowest orders of diffraction are due to the difference in density of protein and water. When we proceed to higher angles, instead of the maxima and minima rapidly dying away, as they would for uniform distribution, they continue quite large. This is of course because we are now finding the effect of the individual points, the outer part of the transform being due to the actual arrangement of these incidents in the structure. In the same way the higher orders of diffraction by the protein are due to the structure inside the molecule. Now comes a very essential point in interpreting a transform, which I should like to stress. You will note that in fig. 5, although the inner maxima are due to the gross structure and the outer maxima to the fine structure of the molecule, what we might call the wave number of the transform, or the rate at which maxima and minima succeed each other, is the same everywhere. It is determined by the overall size of the molecule. If the molecule is large, the maxima and minima, or nodes and loops in the transform, succeed each other rapidly. If the molecule is small they are spaced far apart. The reason for this is obvious. The finest detail in the transform is due to fringes formed by pairs of scattering points at the outer boundaries of the molecule, and the spacing of these fringes is the same as the spacing of maxima and minima due to a uniform distribution of corresponding size.

We can now apply these ideas to the transform of the haemoglobin molecule. There is a property of the haemoglobin crystal which is very advantageous and, indeed, is the basis of the attack on its transform. As was originally shown by

Perutz (1946), it has a number of different forms of shrinkage or expansion due to water entering or leaving the unit cell. Figures 6(*a*) and (*b*) show how this takes place. The molecules remain fixed relatively to each other in layers such as AB A′B′. The distance AA′ also remains constant, but the layers move apart or pack together by an alteration of the angle BAA′. Six different forms assumed by the unit cell are shown in fig. 6(*b*).

If we now think what this means as regards the transform we must remember that the observed amplitudes of the x-ray diffractions are samples of the transform at corresponding points. Since AB remains constant, all these points lie on the layer lines shown in fig. 6(*c*). These are the lines where the path differences for waves scattered by A and B are 0 wavelengths, 2 wavelengths, 4 wavelengths

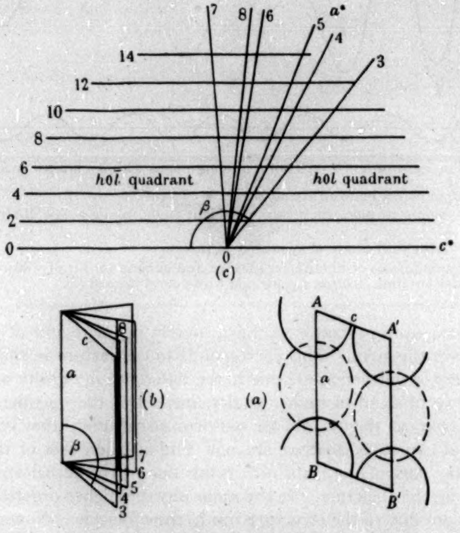


Fig. 6. Exploration of the transform of the haemoglobin molecule by measuring successive stages of shrinkage or expansion of the crystal.

and so forth. Only even numbers appear, because the unit cell has a molecule at the centre of the *c* face, shown dotted in the figure, which has the effect of halving the AB spacing. As the crystal form alters, however, we get many samples of the transform along these lines. The points all refer to the same transform because the molecule retains its orientation with reference to AB.

It is this feature which gives us a chance to interpret the transform. Since the molecule is centro-symmetrical, the values of the transform which are denoted by *F*, the scattering power in the corresponding direction, are either positive or negative. The transform is a series of positive or negative loops separated by zero nodes like those of a vibrating plate. If we have sufficient samples we can try to draw these nodes and loops.

§6.

Figure 7 shows how these nodes and loops can be traced in the case of the reflections by the *c* face of the crystal. The corresponding points appear along the line c^* in fig. 6(*c*). Since initially we do not know their sign, we draw the measured values as lines above and below the origin. The big central maximum must of course be positive, and it will be clear from the figure that it must cross down to B, come up again to C, and go down again to D. It is not clear from the points alone whether the current crosses the origin up to E or whether it dips down again to a corresponding point below the abscissa. We can settle its route by invoking the principle of minimal wave number of the transform which is determined by the size of the molecule, which is known. This wave number is shown by the dotted line below. It is clear that it must cross between D and E, otherwise there would be two maxima too close together for a molecule which we know to be less than 50 Å across. We therefore know that B is negative, C is positive, D is negative, and E is positive, and this is the complete information required to determine a cross section of the molecule in the c^* direction, i.e. perpendicular to the *ab* plane of the crystal.

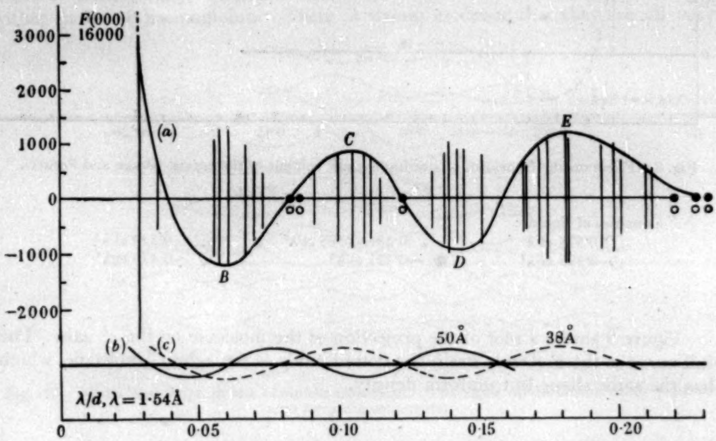


Fig. 7. Nodes and loops along the c^* direction in the transform of haemoglobin.

It is interesting to compare this curve with the one which is obtained when salt is substituted for water. In fig. 8 the dotted curve is for water, the full curve for salt. As is to be anticipated, the salt is diminishing the low-order reflections which are due to the contrast in density of the crystal and its surroundings, while leaving untouched high orders which arise from the intramolecular structure. The difference between these two curves can be shown to represent the scattering of a uniform body which is the same shape as the molecule but has no internal structure.

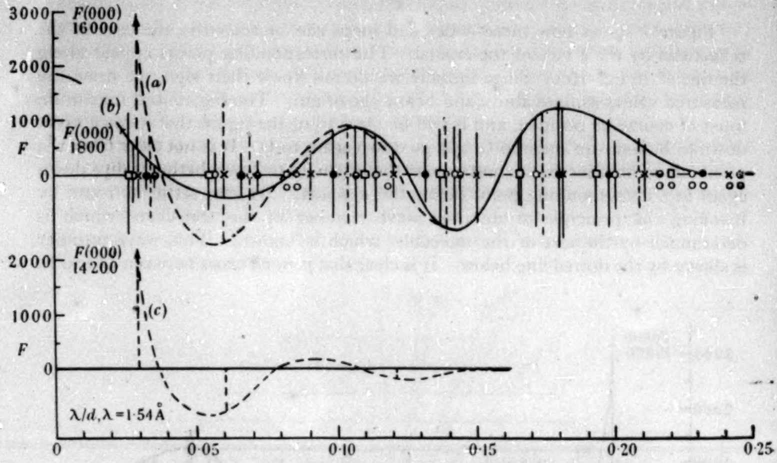


Fig. 8. Effect on the transform of altering the salt content of the crystal (Bragg and Perutz).

- - - - F (salt free) —— F (salt)

Densities of liquids:

⊡ 0·375 el/Å³ × 0·334, 0·395 el/Å³ ○ 0·334 el/Å³

⊙ 0·395 el/Å³ ● ~0·425 el/Å³ ◒ ~0·430 el/Å³

Figure 9 shows a plot of the projection of the molecule on the c^* axis. The full curve is the molecule itself, the dotted curve is the 'ghost' molecule, which has the same shape but uniform density.

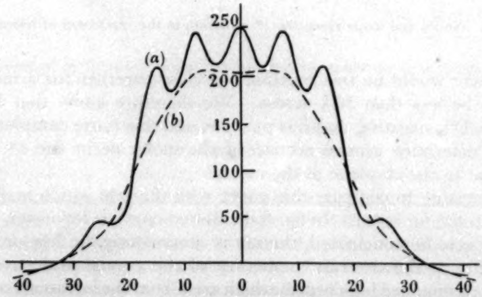


Fig. 9. Projection of the haemoglobin molecule on the c^* axis.

8000 el —— Protein molecule—salt free

7000 el - - - - Salt free—salt (0·395 el/Å³)

§ 7.

The information yielded by fig. 10 alone is of relatively minor importance. What we should like to have is knowledge about the signs over the whole transform, so that we could make a projection of the molecule on the *ac* plane. This involves determining the nodes and loops all over the transform, so that we get the + and − alternations of the F values, and then relating the signs along one layer to those in the next.

We have managed to accomplish the first part of this task. There are still some dubious points where it is not clear whether the F values cross from + to − or remain of the same sign, but it is hoped to clear these up by making more observations. No measurements on this particular crystal, however, can tell us the relations of the signs in the successive layer lines. Since AB remains constant in fig. 6 we cannot get any intermediate points between the layer lines from this form of haemoglobin. To get round this difficulty we have recourse to examining other forms of haemoglobin. There is strong evidence that they are all very closely alike whatever animal they come from. Different crystalline forms will, we hope, establish values of the transform at points between the layer lines, and so enable us to draw nodes and loops up the ladder as well as along its rungs. Now we think we have enough evidence to relate the first two layer lines to the zero one. Figure 10 shows our interpretation of the nodes and loops. The signs given to the nodes, however, are purely speculative; they are merely put there to indicate the kind of thing we hope to determine.

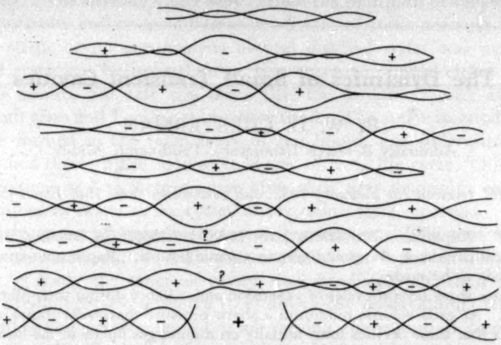


Fig. 10. Nodes and loops in the complete transform. The signs of the layers other than the zero layer are conjectural.

I wish I could have reported to you a solution of the problem and shown you a picture of haemoglobin. I hope, however, that this account of the success which has been achieved so far will illustrate the method of attack. To analyse so complex a crystal would appear at first sight to be an almost impossible problem for X-ray analysis. Fortunately the nature of the heaemoglobin crystal helps us. The swelling and shrinking property and the many forms in which haemoglobin crystallizes provide a wealth of information which is a new phenomenon in X-ray analysis; we have nothing like it in simpler compounds. I feel optimistic about the possibility of a solution. It must be realized, of course, that even if a picture such as we are aiming at here is made, it only represents, as it were, haemoglobin under a low-power objective. We have only tried to establish the transform to a limit which represents detail on about the 7 Å scale. If we can turn this corner, however, I feel that the most difficult pitch has been climbed. I have always been interested in applying x-ray analysis to ever more complex structures, and it would indeed be interesting if we could extend it to the constituents of living matter.

REFERENCES

Bernal, J. D., and Crowfoot, D., 1934, *Nature, Lond.*, **129,** 277.
Boyes-Watson, J., Davidson, E., and Perutz, M. F., 1947, *Proc. Roy. Soc.* A, **191,** 83.
Bragg, W. L., and Perutz, M. F., 1952, *Proc. Roy. Soc.* A, **213,** 425.
Cochran, W., and Crick, F. H. C., 1952, *Nature, Lond.*, **169,** 234.
Pauling, L., Corey, R. B., and Branson, H. R., 1951, *Proc. Acad. Sci., Wash.*, **37,** 205.
Perutz, M. F., 1946, *Trans. Faraday Soc.*, **46B,** 187; 1949, *Proc. Roy. Soc.* A, **195,** 474; 1951, *Nature, Lond.*, **167,** 1053.

The Dynamics of Small Transient Cavities

By W. D. CHESTERMAN
Admiralty Research Laboratory, Teddington, Middx.

MS. received 7th February 1952, *and in amended form* 16*th June* 1952

ABSTRACT. Experimental methods have been developed for creating and observing single spherical cavities in water and certain organic liquids. Rapid photography permits motion analysis to be made.

The cavities show several cycles of expansion and collapse during their lifetime of a few milliseconds. At each volume minimum a sharp pressure maximum occurs in the liquid.

It appears that these cavities form initially on minute gas nuclei in the liquids, and the existence of these nuclei is essential for the type of cavitation studied. The relation of cavity life to maximum diameter is discussed.

§ 1. INTRODUCTION

The problem of avoiding cavitation at the propellers of ships and in hydraulic machinery has always been a most difficult one. Two centuries ago, Euler (1754) referred to the possibility of cavitation whilst developing his theory of turbines. Reynolds (1894) reported the earliest laboratory experiments, and the marine architects at the beginning of this century noted many instances of the damage to naval and other craft (Thornycroft and Barnaby 1895, Silberrad 1912, Ramsay 1912, Parsons and Cook 1919). Fisher (1951) has studied the phenomenon at sea by direct observation.

In the laboratory the complexity of the problem can be studied under more controlled conditions: Rasmussen (1949) used a simple form of water tunnel. The importance of flow conditions in this type of experiment has been emphasized by Konstantinov (1946) and more recently by D. R. Glegg (private communication).

Theoretical studies of cavity dynamics have added only elaboration to the classical paper of Rayleigh (1917). Plesset (1949) used recent experimental cavity studies, and Noltingk and Neppiras (1950) considered other theoretical aspects. Single transient cavities have been created and studied by Briggs *et al.* (1947), Kornfeld and Suvarov (1944), Osborne (1947) and Knapp and Hollander (1948).

Complicating factors are the role of dissolved and 'free' air (Numachi and Kurokawa 1939) and the surface condition of the material near which the cavities form. The closely related problem of the behaviour of water under hydrostatic tension has been studied by Temperley (1946, 1947). The work of Davies and Taylor (1950) on the mechanics of rising bubbles in liquids is also relevant to the cavitation problem. These workers studied the form of bubbles rather larger than the cavities investigated in the present paper.

§ 2. SCOPE OF THE PRESENT STUDY

The object of the research described in this paper was to produce single isolated transient cavities in the body of a liquid and then to study their motion experimentally. The investigations were carried out in liquids of varying surface tension and viscosity and under different conditions of static pressure head.

In these single cavity experiments normal distilled water was used in the experimental systems. No special steps were taken to remove either dissolved or 'free' gas masses, but care was taken that the containers were as clean as possible. Both glass and Perspex tubes were used. The single isolated spherical cavities were formed in the main body of the liquid, and individual bubbles could be studied through the successive stages of their life cycle. The cavities showed oscillations and re-formed again after their first minimum volume; in certain conditions as many as six cycles of cavitation were recorded.

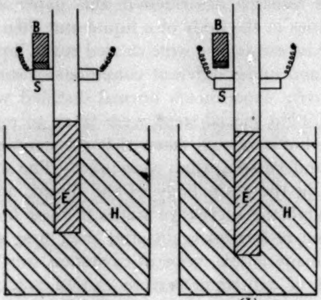
Three main series of experiments were carried out. The first series was carried out with water, and the static pressure head was varied on the cavitation tube surface to study the range of cavity sizes produced and their life cycle. These were preliminary experiments to establish the observational techniques. In the second series, acetone solutions of varying concentration were used in order to examine the effect of surface tension. In the third series a method of achieving single cavity creation was established and the motion of a number of spherical cavities was analysed in detail. The role of free nuclei in controlling the formation of cavities was greatly clarified and the earlier results became much easier to understand. The observational methods were specially developed for the task, and were devised to make the subsequent analysis of the motion as direct and simple as possible. The method of creation of the cavity was simple —the observations were made by taking a very rapid series of short-exposure photographs.

§ 3. EXPERIMENTAL METHOD

(i) *Controlled Cavity Formation*

Various laboratory methods were investigated for producing the cavities in a controlled manner and in a form which could be observed reasonably easily and recorded with precision. One possibility is a body rapidly moving along a predetermined path in water; this produces a 'cloud' of cavities rather than a few which can be isolated, and the turbulent flow created in the wake of the body complicates the physical conditions of the collapse. The alternative arrangement of a stationary body and a controlled fast flow is subject to the same criticism. In the case of flow in a Venturi tube with a constriction, it is easier to reach steady-state conditions, but even then it is difficult to produce and isolate individual cavities. Single cavities can be produced in de-aerated water by the sudden heating of a submerged wire bridge by an electrical current overload. This method is, however, unsuitable for water containing air, because a stream of individual bubbles tends to be formed from a region near the heated wire.

The method finally chosen was of extreme simplicity. A glass tube 1·15 cm in diameter and 100 cm long was half filled with water and mounted vertically in such a way that it could be moved vertically through about 5 cm. Normally the tube is at rest against a rubber buffer as shown in fig. 1 (*a*). It can be pulled down (fig. 1 (*b*)) against rubber-tensioned supports and then suddenly released so that it rises through 5 cm vertically and is then arrested by the rubber buffer.



(*a*) (*b*)

Fig. 1. S. Perspex stop fixed rigidly to moving tube ; B. Fixed metal block with rubber buffer ; H. Submerged hydrophone ; E. Experimental tube.

The water in the tube tends to continue its vertical motion, thus causing a cavity to form near the bottom of the tube when the tube position is as in (*a*). The velocity of the tube at the instant of impact could be controlled and varied from 2 to 6 metres/second. The tube was immersed in an auxiliary optical cell of water for clear distortionless observation. In some experiments a square section Perspex tube of internal section 1·5 cm × 1·5 cm was used, and this had certain advantages in overcoming optical distortion errors in analysis. The mounting for the tubes was made as rigid as possible.

It will be realized that the type of pressure change which is created at the tube base in the body of the liquid is controlled by the force of the shock when the Perspex stop S strikes the metal mount B. The velocity with which this blow occurs is determined by the distance to which the elastic supports are extended and this velocity was rarely increased above 2 metres/second. The suddenness of the blow is defined by the retardation on hitting the block B.

(ii) *Observational Systems*

One of the major experimental difficulties was to devise a method of observation by which the behaviour of the single cavities should be recorded only during the actual period (about 50 to 100 milliseconds) for which the cavities existed, but which would enable about 100 photographs to be taken during this period, i.e. at a rate of between 1 000 and 2 000 pictures per second. It was essential that the images should be recorded on a short length of film and that they should be of the highest quality.

For most of the work a simple drum camera of diameter 15·25 cm was used at a rotation rate of 750 rev/min, giving a film transport speed of 600 cm/sec ; the 48 cm of film wound on the circumference permits 80 pictures to be taken at a picture height of 0·6 cm each. The illuminant was an intermittent flashing light source with very short time of flash: for instance, in 5 microseconds the film (moving at 600 cm/sec) travels 0·030 mm or just over 0·001 inch. This is just within the grain limit of a fast emulsion, so this exposure time set the upper permitted limit of movement during exposure. A power stroboscope was made with the necessary characteristics using a special xenon discharge tube (Chesterman *et al.* 1951 a); the flashing rate was 1000/2000 pictures per second.

There are other advantages in using continuously moving film. A calibrated timing wave can be recorded from a cathode-ray oscillograph for each experimental run, and simultaneous records can be made of pressure changes in the liquid near the cavity (§6). It is possible to make many changes in the physical conditions of the experiment, since each record is on a separate short film.

For some purposes, where the 80 millisecond duration of one revolution of the drum camera is not sufficiently long, a higher capacity camera is useful, giving a time of study of about 1 second (Chesterman and Myers 1951). Picture height in this instrument is that of standard 35 mm film, 1·9 cm at a repetition rate of 2 000 per second or 0·95 cm at a repetition rate of 4 000 per second. The instrument is designed to be linked to the same power stroboscope, in order to maintain the highly desirable feature of short exposure time. As will be seen later, this is especially important in study of the very rapid collapse phases of a cavity.

§ 4. EXPERIMENTAL RESULTS

(i) *Preliminary Work in Water*

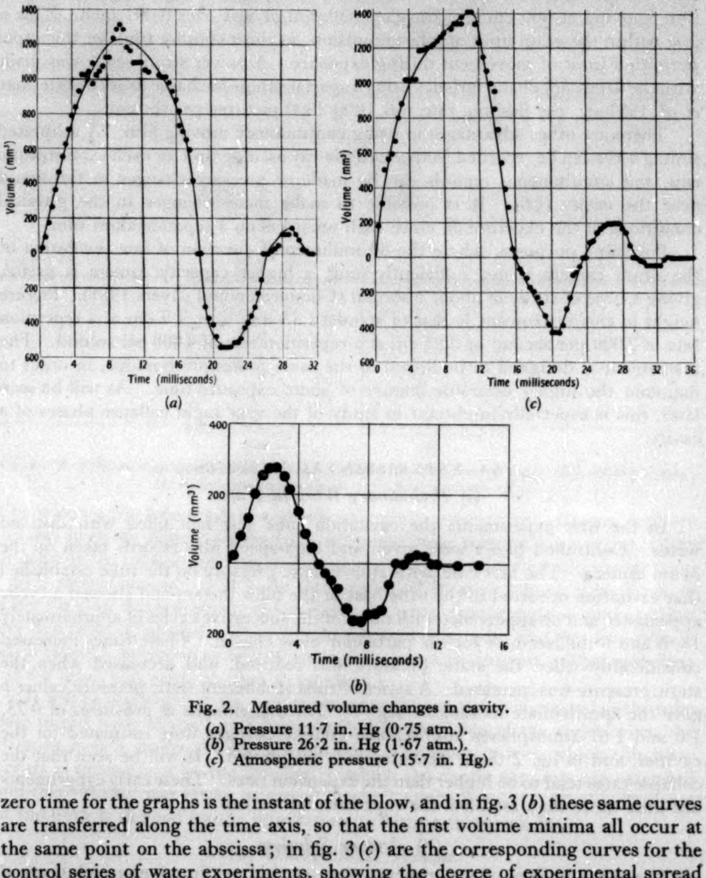
In the first experiments the cavitation tube was half filled with distilled water. Controlled blows were given and high-speed film records taken on the drum camera. The first runs with atmospheric pressure on the tube established that cavitation occurred in the tube base at the blow instant and showed a cyclic appearance and disappearance with times of the successive cycles of approximately 18, 8 and 6 milliseconds for the particular blow chosen. These times increased considerably when the static pressure was reduced, and decreased when the static pressure was increased. A series of runs at different static pressure values p gave the approximate relation $t = Kp^{-1\cdot8}$. For experiments at pressures of 0·75, 1·0 and 1·67 atmospheres approximate volume changes were estimated for the cavities, and in fig. 2 these measurements are plotted. It will be seen that the collapse rates tend to be higher than the expansion rates. These early experiments did little more than establish the experimental method.

(ii) *Acetone Solutions*

The second series of experiments were carried out in aqueous solutions of acetone. The viscosity of pure acetone is 0·303 poises at 19° c, about one-third that of pure water. Surface tension values at 0, 5, 10, 25, 50 and 100 per cent concentrations are respectively 72·0, 55·5, 48·9, 38·3, 30·4 and 23·0 dyn/cm at 25° c.

A set of four control runs was first made with distilled water, and subsequently four runs were made with each of the different concentration solutions. The times of successive oscillation cycles were then measured on the film strips and these are listed in table 1. The interesting fact emerged that the times of the cavitation cycles are almost identical regardless of concentration.

From each group of four records, a complete volume analysis as a function of time was made of a typical cavity. The results are plotted in fig. 3 (*a*) where zero time for the graphs is the instant of the blow, and in fig. 3 (*b*) these same curves are transferred along the time axis, so that the first volume minima all occur at the same point on the abscissa; in fig. 3 (*c*) are the corresponding curves for the control series of water experiments, showing the degree of experimental spread obtained in different runs. Table 2 gives the maximum volumes of the cavities in successive cycles; it will be seen that these tend to increase as the concentration of solution increases.



(*a*) (*c*)

(*b*)

Fig. 2. Measured volume changes in cavity.

(*a*) Pressure 11·7 in. Hg (0·75 atm.).
(*b*) Pressure 26·2 in. Hg (1·67 atm.).
(*c*) Atmospheric pressure (15·7 in. Hg).

Table 1. Acetone Solutions of Different Concentrations.
Times of successive oscillation cycles

Volume (%) of acetone	Strip No.	Times of different cycles (msec) 1st	2nd	3rd	Mean times in cycles (msec) 1st	2nd	3rd
0	39	21·5	8·5	7·0			
	40	25·0	9·5	6·5			
	41	23·0	8·5	—			
	42	22·0	8·5	4·5	23·0	8·7	6·0
5	15*	23·0	9·5	5·0			
	16	26·0	11·5	4·0			
	19	25·5	9·0	4·5			
	20	26·0	10·0	5·0	25·0	10·0	4·5
10	23	25·5	9·0	6·0			
	25	21·5	9·0	6·0			
	26	26·0	9·5	5·5			
	27*	21·5	9·5	7·0	23·5	9·0	6·0
25	28	19·0	11·0	—			
	30	—	10·0	4·0			
	31*	26·5	10·0	4·5			
	32	24·0	11·0	5·5	23·3	10·5	4·5
50	34	—	9·5	—			
	35	23·0	10·0	5·0			
	36*	23·5	10·0	4·5	23·2	10·0	4·7
100	failed to get cavitation						

* These strips were measured for volume changes of the cavity.

Tests were also carried out on carbon tetrachloride. This has the same viscosity as water but a much lower surface tension (26·8 dyn/cm). In three successive runs the times of first cycles were 34·5, 33·5, 34·5 milliseconds, the mean figure of 34 milliseconds being about 50% higher than the value (23 milliseconds) obtained with acetone solutions. It was very difficult to produce cavitation in carbon tetrachloride.

The high speed photographs of these cavities showed an increasing 'diffuseness' in the cavity appearance at each successive cycle. The breakdown of the cavity form at collapse is violent and most rapid.

Table 2. Acetone Solutions of Different Concentrations.
Maximum volumes of cavity in successive cycles

Volume (%) of Acetone	Strip No.	Maximum volumes in each cycle (mm^3) 1st	2nd	3rd	4th	Tube vel. at impact (m/sec)
0	39	1150	120	50	—	2·00
5	15	1090	340	100	—	2·48
10	27	1300	140	70	50	2·10
25	31	1330	350	170	120	2·36
50	36	2080	480	290	170	2·50

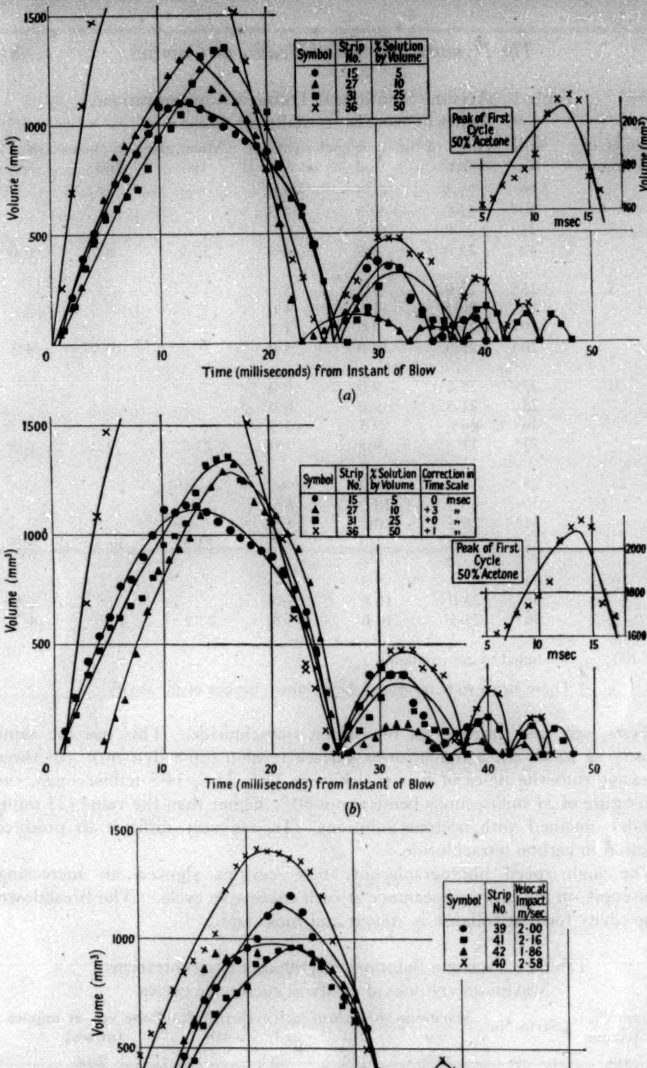


Fig. 3. (*a*) Acetone solutions of different concentrations. Estimated volume changes in cavity. (*b*) Acetone solutions of different concentrations. Estimated volume changes in cavity. Curves plotted so that first volume minima are coincident. (*c*) Control experiments on distilled water. Estimated volume changes in cavity. Curves plotted so that first volume minima are coincident.

(iii) *Influence of Air Nuclei*

Theoretical work on the tensile strength of liquids has been largely confined to fluids which are free from gas nuclei, since the existence of these nuclei greatly reduces the strength of the liquid and also introduces a difficult discontinuity in the treatment of the problem. The theories of Frenkel (1946), Fürth (1941), Döring (1939) and others all predict that surface tension is the only significant property of a pure liquid which distinguishes it from another as regards its tensile strength, but it is clear that in the experimental conditions with which we are dealing this consideration is not relevant. Thus, in cavity formation it is not only the dissolved gas that is of importance but also the minute gas bubbles or 'air micro-nuclei'. These results provide additional evidence for the ideas of Harvey *et al.* (1947) who obtained the remarkable result that if water is carefully treated to remove completely all gas nuclei then a blunt-ended glass rod can be fired through the liquid at 37 m/sec without cavitation, although at that speed the liquid near the end of the rod must be in tension. Pease and Blinks (1944) have shown that when these nuclei are not removed, they can and do lead to bubble formation. Extreme experimental precautions are necessary if it is desired to study cavitation in water in the absence of any pre-existing gas phase. In the absence of any free gas phase, rupture of the liquid itself only occurs under very high tensile forces. There is no 'focus' in the liquid to allow easy expansion and consequent relief of the tension in the water. For instance water in this condition will not boil until above 200° c and the boiling then occurs explosively.

In the third series of shock-created cavity experiments there is strong evidence that it is the air nuclei which expand to form the initial cavities. The cavities observed are formed easily at low pressure differences and, under carefully controlled experimental conditions, expand as isolated spherical cavities.

In the acetone solutions, there are considerable changes in surface tension as the concentration increases, and in addition the viscosity decreases with increase of concentration. Since both variables change during the series it is difficult to assign the relative importance of the two physical characteristics but the results quoted for carbon tetrachloride show that decrease of surface tension tends to increase the cycle times and that surface tension has a more marked effect as the cavity size becomes very small. In the acetone series, therefore, there are the two opposing parameters of decreasing viscosity, tending to make the cycles less rapid. It is felt, however, that neither of these physical factors are of dominant importance in controlling the behaviour of these cavities, or indeed the volume changes in the different series to any great extent. Acetone differs from water in certain other respects, and it may be that these other differences are more important.

As the concentration of the acetone solutions increases it was found experimentally that more vigorous 'priming' of the liquid was necessary before cavitation occurred: repeated blows were necessary with the tube of the 50% solution before cavities formed. Carbon tetrachloride required similar treatment. It is probable that by this treatment gas nuclei are formed by forcing air out of solution, so that foci for cavitation then exist. The blow which was filmed and analysed took placed immediately after the 'priming' before the nuclei had gone back into solution. In water these are knocked out as bubbles. Such an explanation as this would account partly for the great similarity of results on all the different solutions in regard to cycle times and cavity volume changes.

In the distilled water experiments carried out, the diameters of the 'free' spherical air nuclei are certainly below 100 microns, so that their volume is of the order of 5×10^{-4} mm^3; even at a population density as high as 20 per cm^3 of liquid, their combined volume is therefore less than 0·1 mm^3, a figure small compared with the value of 6 mm^3 total oxygen content per cm^3 in normal water. These nuclei are therefore only revealed by the cavitation experiment itself. Tap water may show the same dissolved gas on a Winkler oxygen test when fresh and after standing one month, but its ability to cavitate may be entirely different.

(iv) *Single Spherical Cavities*

At this stage in the research it was clear that the formation technique had to be more carefully controlled if it were desired to study the life cycle of a single spherical cavity. Rare occurrences of single cavities were too infrequent to provide useful experimental data. In the observational system used the cavities were photographed in the first two series of experiments at about 20 diameters reduction on the negative, so that the tube end gave an appearance of being filled with a cavity and the fine structure would not be seen; it became apparent that close-up studies of about 2 diameters reduction would be necessary. Both these changes were made in the third series of experiments. In addition, by suitable choice of two independent variables—number and type of pre-treatment blows—and type of final blow various types of cavity cluster could be systematically created.

As a qualitative statement the pre treatment 'priming' blows may be graded as 'Severe', 'Medium' and 'Weak' and the same terminology may be applied to the final blow itself which formed the cavities recorded photographically. Different numbers of treatment blows were given, and this pre-treatment and the final blow determined the number and maximum size of the cavities formed. For instance, when the tube is strongly 'primed' (perhaps by 10 severe blows) then a large cluster of many cavities appears in the tube. With rather weaker priming a small cavity cluster of about 50 spherical cavities occurs. With no 'priming' there are often formations on the Perspex wall surface and not in the body of the liquid.

As in the earlier experiments there were several cycles in the life history of the cavity clusters. Individual spherical cavities in these small clusters are sufficiently separated from each other to permit their dynamics to be studied.

§ 5. DISCUSSION OF RESULTS

(i) *Life and Collapse Times*

Consider in more detail the individual cavities in the first cycle of fig. 4 (*a*), a cluster of 50 spherical cavities in a volume of 3 cm^3 of water. It is difficult to believe that these are formed so rapidly in this way unless they expand on existing nuclei during the first millisecond. It will be seen that the largest cavity has the longest life, the small cluster in the right-hand top corner of the picture vanishing at frame 3, and the whole field is clear 7 milliseconds after the blow.

In order to investigate the dependence of collapse time on maximum diameter of the cavity, about six individual cavities were studied on each of four different typical records of the sort shown in fig. 4. Analysis was made on a total of 23 cavities, of which 18 were spherical in form, four formed on a surface and one was irregular in shape. The variations of diameter with time are plotted for these cavity groups in fig. 5. From these curves the maximum diameters reached may be estimated.

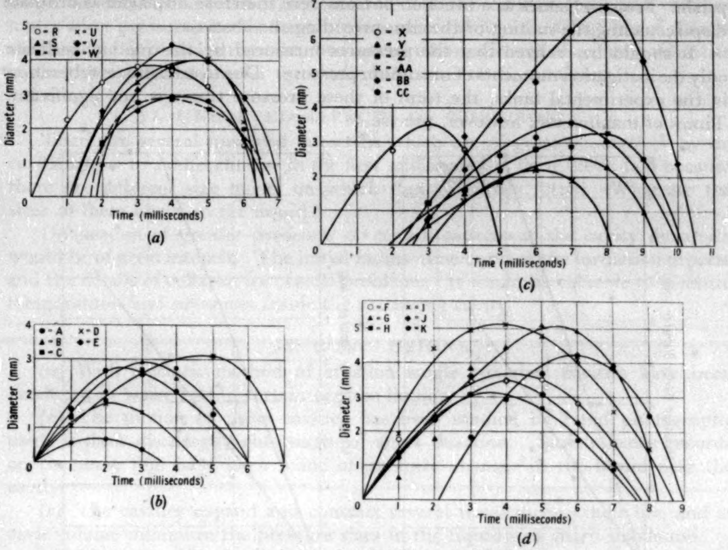


Fig. 5. Variation of cavity diameter with time.

In fig. 6 maximum diameter is plotted against collapse time for these cavities. It will be seen that there is a considerable scatter of points, but that these cavities all collapse at about half the speed predicted by the Rayleigh formula. Figure 6 also shows the variation of life with maximum diameter. The steeper of the two lines drawn assumes a life double the Rayleigh collapse time.

(ii) *Interaction Effects*

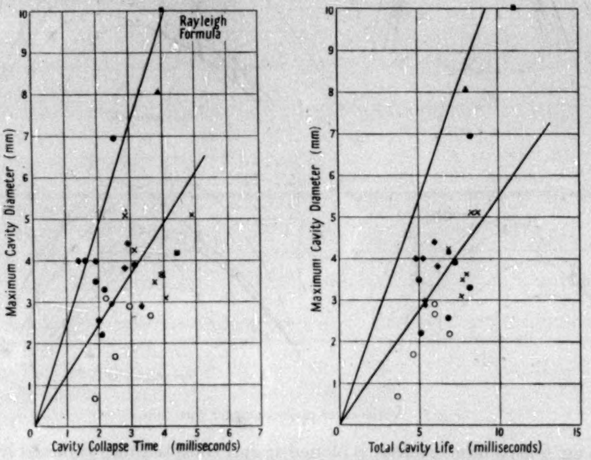
Figure 4 (*b*) shows what happens when expansion of a cavity takes place into an adjacent cavity. Whilst the cavity in the top left-hand corner has an undisturbed life, the three middle cavities all move into the lower cavity and form a larger cluster, which then collapses. Figure 4 (*c*) shows the formation and collapse of a single hemispherical cavity on the solid Perspex surface. This cavity was the result of a very gentle pressure change.

§ 6. PRESSURE MEASUREMENTS NEAR THE CAVITIES

(i) *Method of Measurements*

Throughout all these experiments the pressure changes in the liquid near the cavities were recorded on the same drum camera film on which the growth and collapse pictures were taken. The procedure used was to insert a rochelle salt hydrophone in the liquid at a distance of 5 cm from the cavitation tube in the main surrounding water tank. The hydrophone output was amplified and passed to a high-brightness cathode-ray oscillograph. The spot displacements were directly recorded by an auxiliary optical system of the camera using a prism. Pressure values as a function of time were therefore displayed as ordinate displacements, the motion of the film providing the abscissa.

It should be realized that the pressures measured by this method indicate only the instigation moments of maximum pressure. Due to acoustic reverberation in the experimental tanks, the form of these pressure traces is not significant. Times of maxima are, however, precise.



(*a*) (*b*)

Fig. 6. (*a*) Dependence of collapse time on maximum cavity diameter ; (*b*) Dependence of total cavity life on maximum diameter.

(ii) *Incidence Time of the Pressure Maximum*

The important fact emerged that the times of minimum cavity volume were those of maximum pressure recorded at the hydrophone. This is a result which might have been expected from explosion bubble theory, but clear experimental proof did not exist of this until these records were made. This result has been obtained many scores of times on records in various liquids. With this type of cavity, no pressure is developed during the initial formation cycle of the cavity, but at the collapse moment of each cycle these pressures rise rapidly to maxima. The pressure peak is reached in the order of 5 to 10 microseconds. It is this pulse which causes erosion damage.

(iii) *Instantaneous Hydrodynamic Pressures*

There remains the vital question as to the nature of the hydrodynamic pressure field under which the cavities grow and collapse. A valuable contribution to the solution to this problem has been achieved recently by the author's colleague

D. R. Glegg (private communication) who used a different form of pressure measurement. He employed a circular experimental tube whose base was the disc of a condenser-type hydrophone, and he formed the cavities on this surface plate. Recording the instantaneous output of this pressure gauge on the oscillograph he obtained records which showed that the cavities grow and collapse under a low pressure near the vapour pressure of the liquid. Glegg has explained the form of the pressure trace he obtained by taking account of acoustic velocities in the tubes and allowing for the elastic behaviour of the container.

§7. UNSOLVED ASPECTS OF THE PROBLEM

There are several unsolved questions which are of great interest. Why do cavities grow to a different size in the first millisecond of their life? Is it because there are different size nuclei on which they originally form? What are the sizes of these nuclei in the liquid?

Delineation of greater precision of certain aspects of the cavity dynamics would be of great interest. The initial radius–time curve at the formation process and the details of collapse are crucial problems. It would be valuable to measure temperatures and pressures inside the collapsing cavity.

§8. CONCLUSIONS

(*a*) With a shock method of creation single spherical cavities have been produced in water and in certain organic liquids.

(*b*) The motion of these cavities has been studied by rapid photography using a flash discharge light source of short duration. Simultaneous records on the same film have been made of pressure changes in the liquid near the cavity.

(*c*) The cavities expand and contract several times during their life, and at each volume minimum the pressure rises in the liquid to a sharp maximum.

(*d*) Life and collapse time of these cavities are directly related to the maximum diameter to which the cavity grows. Satisfactory agreement was achieved with Rayleigh's theoretical predictions.

(*e*) The experiments strongly support the concept that gas nuclei in the liquids act as centres for cavity formation. As other workers have found, without such nuclei it is difficult to produce cavitation under these stresses.

(*f*) Examination of liquid mixtures having different surface tensions did not lead to precise results. This is partly due to viscosity changes and partly due to the overriding effect of variations of size and stability of the gas nuclei.

ACKNOWLEDGMENTS

It is a great pleasure to acknowledge the help given by my colleague D. R. Glegg (Admiralty Research Laboratory) throughout this work. Dr. A. B. Wood was kind enough to make various suggestions about the manuscript.

The paper is published with the permission of the Admiralty.

REFERENCES

Briggs, H. B., Johnson, J. B., and Mason, W. P., 1947, *J. Acoust. Soc. Amer.*, **19**, 664.

Chesterman, W. D., Glegg, D. R., Meadowcroft, A. J., and Peck, G. T., 1951, *Proc. Inst. Elec. Engrs.*, **98**, 619.

Chesterman, W. D., and Myers, D. T., 1951, *J. Sci. Instrum.*, **28**, 301.

Davies, R. M., and Taylor, C. I., 1950, *Proc. Roy. Soc.* A, **200**, 375.

Döring, W., 1939, *Z. Phys. Chem.*, B, **38**, 292.
Euler, L., 1754, *Hist. Acad. Roy. Science et Belles Lettres*, **10**, 221.
Fisher, J. N., 1951, *Engineer, Lond.*, **192**, 656.
Frenkel, J., 1946, *Kinetic Theory of Liquids* (Oxford : Clarendon Press).
Fürth, R., 1941, *Proc. Camb. Phil. Soc.*, **37**, 252.
Harvey, E. N., McElvoy, W. M., and Whiteley, A. M., 1947, *J. Appl. Phys.*, **18**, 162.
Knapp, R. T., and Hollander, A., 1948, *Trans. Amer. Soc. Mech. Engrs.*, **70**, 419.
Konstantinov, V. A., 1946, *Izv. Akad. Nauk, S.S.S.R., Otdel. Tekn.*, **10**, 1355.
Kornfeld, M., and Suvarov, L., 1944, *J. Phys., U.S.S.R.*, **8**, 171.
Noltingk, B. E., and Neppiras, E. A., 1950, *Proc. Phys. Soc.* B, **63**, 674.
Numachi, F., and Kurokawa, T., 1939, *Werft-Reederei-Hafen*, Vol. 20.
Osborne, M. F. M., 1947, *Trans. Amer. Soc. Mech. Engrs.*, **4**, 253.
Parsons, C. A., and Cook, S. S., 1919, *Engineering*, **107**, 501, 515.
Pease, D. C., and Blinks, S., 1944, *J. Cell. Comp. Phys.*, **24**.
Plesset, M. S., 1949, *J. Appl. Mech.*, **71**, 277.
Ramsay, W., 1912, *Engineering*, **100**, 687.
Rasmussen, R. E. H., 1949, *Trans. Dan. Acad. Sci.*, **1**, 1.
Rayleigh, Lord, 1917, *Phil. Mag.*, **34**, 94.
Reynolds, O., 1894, *Brit. Ass. Adv. Sci. Report* 564, Science Papers 2.578.
Silberrad, D., 1912, *Engineering*, **100**, 34.
Temperley, H. N. V., 1946, *Proc. Phys. Soc.*, **58**, 420, 438; 1947, *Ibid.*, **59**, 199.
Thornycroft, J., and Barnaby, S., 1895, *Proc. Inst. C. W. Engrs.*, **122**, 51.

The Temperature Variation of the Magnetization of Nickel in Low and Moderate Fields

By R. S. TEBBLE, J. E. WOOD* and J. J. FLORENTIN
Physics Department, University of Leeds

Communicated by E. C. Stoner; *MS. received 10th June* 1952

ABSTRACT. An account is given of measurements on the reversible changes in magnetization accompanying change in temperature in annealed nickel, and the results are discussed in relation to the work of Bates and others on the temperature changes produced by adiabatic changes in magnetization (the magneto-caloric effect). It is found that at relatively high fields, where the magneto-caloric effect is mainly reversible, the results are in good agreement with those obtained from direct magneto-caloric measurements, while for low field values a quantitative separation of the reversible and irreversible components of the magneto-caloric effect has been carried out. The result of an analysis of the reversible heating effect is compared with the estimate given in the theoretical investigation of Stoner and Rhodes, and it is suggested that, over the low field range considered, rotational processes predominate.

§ 1. INTRODUCTION

In recent years considerable progress has been made in gaining a fuller knowledge of the various elementary processes contributing to the change of magnetization of a ferromagnetic with changing field. It is, however, still far from possible to determine with any certainty the relative contributions made by these various processes over successive parts of an initial magnetization curve or a hysteresis loop for ordinary polycrystalline materials. A partial analysis may be made by determining the reversible susceptibility at points along a magnetization curve; this gives the reversible part of the change, and the difference between the reversible and incremental susceptibility (given by the slope of the magnetization curve) gives the irreversible part. Accurate measurements have now been made of the reversible susceptibility of a number of materials at room temperature (Tebble and Corner 1950), and also, using a new semi-automatic method, of various specimens of nickel at a series of temperatures between about -50°c and $+200^\circ$c (Tebble, Corner and Wood 1951). Much detailed information has also been obtained about the size distribution of the discontinuities which make up the irreversible part of the change in magnetization by studies of the Barkhausen effect (Bush and Tebble 1948, Tebble, Skidmore and Corner 1950). Owing to the limit imposed experimentally on the smallest size of discontinuity which can be detected, the aggregate of the observed discontinuities does not account for the whole of the irreversible change, and extrapolation beyond the experimental range, for individual parts of the magnetization curve, is very uncertain. The total irreversible change can, however, be obtained by difference, as already indicated.

---

* Now with the Royal Naval Scientific Service.

To determine the precise character of the reversible changes taking place much further information is required. An important means of obtaining further information is by the study of the magneto-caloric effect (the temperature changes accompanying adiabatic magnetization), on which very extensive results have been obtained in the last ten years by Bates and his collaborators (for a summarizing survey see Bates 1949, Stoner 1950). A comprehensive theoretical investigation of this effect has been carried out by Stoner and Rhodes (1949) in which the physical processes involved are discussed, and a partial analysis of some of the results is given. It is pointed out, however, that "it is essential for a proper understanding of the magneto-caloric effects that an unambiguous separation of the reversible and irreversible heat changes should be made".

It might be suggested that the obvious method of estimating the reversible component of the magneto-caloric effect would be to carry out the experiments of Bates and Davis (1950) under reversible conditions. It has been shown from measurements on reversible susceptibility (Tebble and Corner 1950) that to produce reversible changes in magnetization in the 'coercivity' region, changes in field as small as 10^{-2} oersted are required. In order to carry out magneto-caloric experiments for correspondingly small field changes it would be necessary to detect temperature changes of the order of 10^{-8} deg. c. As it seems unlikely that there can be any appreciable improvement on the sensitivity achieved by Bates and his co-workers (10^{-6} deg. c), alternative methods must be used to estimate the reversible components of the magneto-caloric effect in the lower field range. Okamura (1936) has attempted to separate the reversible and irreversible components but his methods are not completely satisfactory although his results do give an indication of what might be expected for high field values. Stoner and Rhodes (1949, p. 509) suggested that this separation might be effected through measurements of the reversible susceptibility at different (neighbouring) temperatures, followed by an application of a thermodynamic relation between the adiabatic change of temperature with field and the change with temperature of the reversibly attained magnetization. Further consideration, however, has shown that this argument is defective. The present work involves a completely justifiable application of the relevant thermodynamic relation, which links together a reversible isentropic change of temperature with field and a reversible change of magnetization with temperature at constant field.

In the experiments to be described systematic measurements have been made of the change of magnetization with temperature, at constant field, for specimens of nickel at points along the magnetization curve. For any considerable change of temperature the change in magnetization is not in general reversible, and in cyclic changes complex hysteresis effects may be obtained. By making the temperature change sufficiently small, however, it is possible to make good estimates of the reversible component of the change of magnetization. From the results the reversible component of the magneto-caloric temperature change may be derived, and from its magnitude tentative conclusions may be drawn, following the treatment of Stoner and Rhodes, about the character of the elementary processes contributing to the reversible change in magnetization with field.

If strictly comparable results are to be obtained in the different types of measurements on a given material, ideally the same specimen should be used. This has not been practicable, but the present measurements on nickel have been made on material approximating closely in composition and treatment to that used in the previous work on reversible susceptibility. The magnetic characteristics are sufficiently close to those of the nickel used by Bates and Davis (1950) in their magneto-caloric measurements for a significant comparison between the two sets of results to be made.

§2. THEORETICAL

In the experiments on the magneto-caloric effect the specimen is taken round a hysteresis loop in steps of 1 to 30 oersted, giving temperature changes of the order 10^{-5} deg. C. The results are usually given in the form of (Q', H) curves, where Q' is given by

$$Q' = \sum_{-H_{\max}}^{H} \delta Q' = \sum_{-H_{\max}}^{H} C\rho\,\delta T'. \qquad \ldots\ldots (1)$$

C and ρ are the specific heat and density, respectively, and $\delta T'$ is the change in temperature resulting from an adiabatic change in the applied field H. Thus Q' is the aggregate change in thermal energy developed in traversing the hysteresis loop from the initial state $-H_{\max}$ to H. It is not to be confused with the heat intake dQ in the expression for the change in internal energy $dE = dQ + H\,dI$, where dI is the change in intensity of magnetization. The entropy change dS is related to dQ by $dQ \leqslant T\,dS$, where the equality sign applies to reversible changes only. Since the change in magnetization takes place adiabatically $dQ = 0$, but the entropy change is zero only if the change is reversible. The change in internal energy may be divided into two parts, a magnetic energy change dE_M and a thermal energy change dE_T, so that

$$dE = dE_M + dE_T = H\,dI \quad \text{and} \quad \Delta E_M = \int H\,dI - Q'. \qquad \ldots\ldots (2)$$

The derivation of this and most of the other thermodynamic relations in this paper is given in Stoner and Rhodes (1949). Equation (2) is valid for both reversible and irreversible adiabatic changes, and the experimental results can thus be expressed in the form of a change in internal magnetic energy. Stoner and Rhodes, in their theoretical investigations of the magneto-caloric effect, have developed a method of analysis in which the measured heat change is written

$$Q' = a\int d(IH) + \int b'' H\,dI. \qquad \ldots\ldots (3)$$

The first term represents that part of the heat change associated with a change in intrinsic magnetization I_0 with the constant a given by

$$a = -(T/I_0)\,dI_0/dT. \qquad \ldots\ldots(4)$$

As a can be evaluated with fair accuracy from independent experimental results a knowledge of the heat changes provides a means of estimating the 'experimental' coefficient b'' in (3). For purely reversible changes the second term arises from reversible processes other than change in intrinsic magnetization, these being rotations of the domain magnetization vectors and reversible domain boundary movements. It is shown theoretically that if, of these, only reversible rotational changes occur, and the material is free from strain,

$$Q' - a\int d(IH) = b\int H\,dI, \qquad \ldots\ldots(5)$$

where $b = (T/k)dk/dT$ and k the modulus of the anisotropy coefficient. The results of an analysis of the magneto-caloric measurements by the method of Stoner and Rhodes are summarized in a curve giving b'' as a function of H. The analyses which have been made of results on annealed materials indicate that, for the higher field range, b'' is approximately constant, and that it is in fair agreement with the value of b as estimated for rotational processes, as discussed more fully by Stoner and Rhodes. In the lower field range, there are rapid variations in the value of b'' which are attributed in the main to the complicating effect of the superposition of irreversible and reversible changes. As the usual thermodynamic relations are applicable only to changes which are reversible, the more complete analysis of the magneto-caloric effect requires the separation of the reversible and irreversible components Q'_{rev} and Q'_{irrev} which make up Q', i.e.

$$Q' = Q'_{\text{rev}} + Q'_{\text{irrev}}. \qquad \ldots\ldots(6)$$

An indirect method of estimating Q'_{rev} is founded on the thermodynamic relation for a reversible process,

$$(\partial T/\partial H)_S = -(T/C_p)(\partial I/\partial T)_H, \quad \text{or} \quad (\partial Q'/\partial H)_S = -T(\partial I/\partial T)_H. \quad \ldots\ldots(7)$$

The reversible heat change is then given by

$$Q'_{\text{rev}} = \int(\partial Q'/\partial H)_S\,dH = -\int T(\partial I/\partial T)_H\,dH, \qquad \ldots\ldots(8)$$

integrated between the appropriate limits.

This is the basis of the method described in this paper which enables Q'_{rev} to be estimated from measurements of the change of intensity of magnetization I with temperature, i.e. $(\partial I/\partial T)_H$, under conditions which satisfy the requirements of reversibility. From the results the corresponding values of b'' have been calculated for field values ranging from -50 oersted to $+50$ oersted.

§ 3. METHODS OF MEASURING $(\partial I/\partial T)_H$

In a 'direct' method of determining $(\partial I/\partial T)_H$, by measuring the change in magnetization as the temperature of the specimen is changed, the range over which the change is reversible will vary over the magnetization curve. Experimentally, the obvious test for reversibility is that on increasing and decreasing the temperature of a specimen the magnetization should return through the same intermediate points to its original value. A practically useful statement of a sufficient, but not necessary, condition for reversibility over a given range $\delta T_{\max}$ is that within this range $\delta I/\delta T$ should be independent of δT. It is possible to estimate the order of δT required for reversibility in the hysteresis region from Tebble and Corner's (1950) measurements on reversible susceptibility. With $\delta I = 0{\cdot}2$ e.m.u. and $(\partial I/\partial T)_H = 0{\cdot}2$ the corresponding value for δT would be about 1° c. The problem is thus essentially one of devising a method of measuring small changes in intensity resulting from temperature changes of this order.

In the high field region a value of $(\partial I/\partial T)_H$ can be obtained from a knowledge of the magnetization curves of a specimen over a range of temperatures, and as a preliminary experiment measurements on the magnetization curves of a nickel rod have been made (§4(i)). In the hysteresis region it is necessary to take a specimen step by step through a hysteresis cycle, and at each step to measure the change in magnetization in the specimen resulting from a small change in temperature. This involves the use of a magnetometer to measure the corresponding change in field at a distance from the specimen: the use of a galvanometer with a search coil wound on the specimen is impracticable owing to the extreme difficulty of eliminating thermoelectric effects produced in the search coil as the temperature is changed.

In the experiments of Bates and Davis on the magneto-caloric effect cylindrical nickel rods 40 cm long and 4 mm diameter were used; it would obviously have been an advantage in any comparison of the results if the specimens used in the present work were of similar shape and dimensions, but the flux distribution in such a specimen (and hence the 'effective pole separation') changes over the magnetization cycle in a manner which cannot be easily calculated with any degree of accuracy (Bozorth and Chapin 1942). In the method of measuring changes in magnetization by means of a search coil wound round the centre of the specimen this effect is small but, since it is only possible to use a magnetometer if it is placed at a distance from the rod such that the effect of changes in effective pole separation is small, the sensitivity of the magnetometer to changes in intensity in the rod is thus correspondingly low, giving results which are of significance only in the high field region. For measurement in the coercive region it is necessary to use a specimen in the form of an ellipsoid; it is then simple to calculate the intensity of magnetization from a knowledge of the field at a point outside, and the magnetometer may be placed as close as possible to the specimen. The expressions used in calculating the field due to an ellipsoid are given in the Appendix.

§ 4. DESCRIPTION OF EXPERIMENTS

(i) $(\partial I/\partial T)_H$ *from Measurements on the Magnetization Curves*

The specimen used in these experiments was a nickel rod 40·8 cm long and 0·448 cm in diameter; the composition and magnetic properties of this specimen (A) and of the other specimen used in this work are summarized in table 1. The normal 'ballistic' method was used; a search coil was wound on the centre of the specimen and a Grassot fluxmeter was used to detect the change in flux. The rod and search coil were contained in a glass tube through which the cooling and heating liquid were pumped and this assembly was contained in the magnetizing solenoid. Considerable care was taken in the measurements, and it was possible to reproduce hysteresis loops with errors in I of less than ± 2 e.m.u. The temperature range covered was from $-63°$ c to $+69°$ c, the average interval being about 15° c. From curves of I plotted against T, at constant H, $(\partial I/\partial T)_H$ has been found over a range of temperatures; the results for 20° c and 40° c are of particular interest as they can be compared with the results of Bates and Davis (1950), and Hardy and Quimby (1938), whose measurements on the magneto-caloric effect were made at about 18° c and 41° c respectively (fig. 1).

Table 1

	Nickel A (rod)	Nickel B (rod)	Nickel C (ellipsoid)
Length (cm)	40·8	45·0	43·1
Diameter (cm)	0·488	0·45	0·422
H_c	2·9	0·17	0·36
$I_{\text{remanence}}$	238	99	198
$I_{H=50}$	429	430	424
Composition (%)	Ni 99·5	Ni 99·71, Fe 0·06, Mn 0·01, Si 0·06, C 0·02, Mg 0·13, S 0·008.	

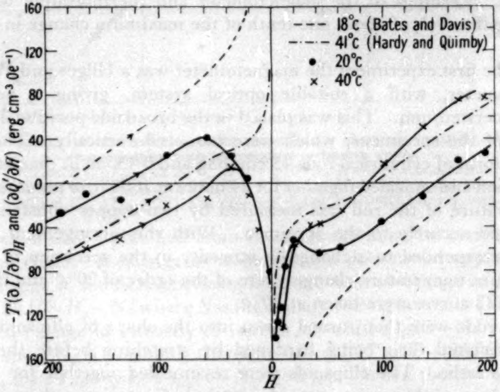


Fig. 1. Temperature variation of the magnetization of annealed nickel A. Full curves $-T(\partial I/\partial T)_H$ from present experiments; broken curves $(\partial Q'/\partial H)_S$ from magneto-caloric measurements.

$-T(\partial I/\partial T)_H$ has been plotted in fig. 1 since, in so far as the changes are reversible, it gives the magneto-caloric quantity $(\partial Q'/\partial H)_S$ (see eqn. (7)).

The results show clearly that $(\partial Q'/\partial H)_S$ is markedly temperature dependent, and emphasize the importance of stating the temperature of the experiment in any work of this nature. (Unless this temperature is given, considerable uncertainty is introduced in the analysis of the results of magneto-caloric experiments, as pointed out by Stoner and Rhodes 1949, pp. 513, 514.) The general agreement between the measurements and those of Bates and Davis and of Hardy and Quimby outside the hysteresis region is reasonable, particularly in view of the differences which must exist between the specimens.

(ii) *Magnetometer Experiments*

The three sets of magnetometer experiments which have been made are basically similar, but in the successive sets the detailed arrangements were modified to give greater sensitivity, so that accurate measurements could be made of the changes of magnetization for smaller changes of temperature. With the most sensitive arrangement reliable determinations could be made of the reversible component of the change of magnetization with temperature over the whole of the hysteresis region. The results of all these experiments are given here because a useful check on the reliability of the results is obtained in those field ranges where there is overlapping.

In each case two specimens were used, mounted within identical magnetizing solenoids and placed equidistant from the magnetometer. The solenoids were connected in series-opposition so as to balance out approximately the resultant field at the magnetometer. With the temperature of one specimen kept constant by a flow of water, the temperature of the other was raised or lowered by a suitably controlled continuous flow of butyl alcohol, heated electrically or cooled in an ice-mixture. Readings of the magnetometer and thermocouple were taken at temperature intervals of about one-tenth of the maximum change in temperature $\delta T_{\max}$.

(*a*) In the first experiment the magnetometer was a Hilger and Watts vertical force variometer, with a suitable optical system, giving a sensitivity of $4{\cdot}80 \times 10^{-5}$ oersted/mm. This was placed in the broadside position 32·2 cm from the centre of the specimens, which were mounted vertically. The specimens were in the form of cylindrical rods 45 cm long and 0·45 cm in diameter, identical in composition and annealed together for 24 hours at 1000° C *in vacuo* (specimen B). The temperature of the rod was measured by two copper–constantan thermocouples wired securely to the specimen. With this arrangement a deflection of 1 mm corresponded to a change in intensity in the specimen of 0·38 e.m.u. The maximum temperature changes were of the order of 20° C and the gradients of the (δI, δT) curves were taken at 20° C.

(*b*) The rods were then turned down into the shape of ellipsoids, a process which necessitated their being hardened by stretching before they could be worked on a lathe. The ellipsoids were re-annealed together for 40 hours at 1000° C *in vacuo*; they were not of identical dimensions, the compensating ellipsoid being slightly larger than that on which the measurements were made (specimen C). A horizontal suspension magnetometer of the Sucksmith type (1945) placed in the end-on position, 8·3 cm from the end of the horizontal specimen, gave a sensitivity of 0·092 e.m.u./mm ($8{\cdot}05 \times 10^{-5}$ oersted/mm), an improvement in sensitivity of about 4 : 1. A single copper–constantan thermocouple fixed to the specimen was used to measure the changes in temperature ($\delta T_{\max} \simeq 5°$ C).

(*c*) The Sucksmith magnetometer was then made astatic by the addition of a second set of magnets 12·63 cm below the original set, and placed 8·43 cm from the end of the specimen. With a suitable controlling field an average sensitivity of 0·0228 e.m.u./mm ($1{\cdot}15 \times 10^{-5}$ oersted/mm for one pair of magnets) was obtained, enabling measurements to be made with a maximum temperature change $\delta T_{\max} \simeq 1{\cdot}3°$ c. Fields of 100 oersted in the solenoids result in fields of the order of 1 oersted at the magnetometer, and, in order to reduce the magnetometer fluctuation to about 0·1 mm over most of the field range, it was necessary to cool the windings of the solenoids by means of a continuous flow of liquid.

§ 5. RESULTS

Typical sets of readings, showing magnetometer deflection ($\propto \delta I$) plotted against thermocouple reading ($\propto \delta T$) are given in fig. 2(*b*) and 2(*c*). An indication of the improvement in the experimental conditions may be obtained from a comparison of fig. 2(*a*), which represents a set of 'irreversible' results obtained for nickel B with $\delta T_{\max} \simeq 20°$ c, and fig. 2(*b*), which shows the results obtained for nickel C at approximately the same point on the hysteresis loop, with $\delta T_{\max} \simeq 1{\cdot}3°$ c.

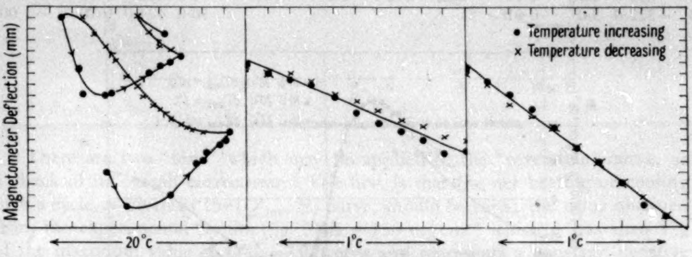


Fig. 2. Typical results of the magnetometer experiments showing variation of intensity of magnetization with temperature (arbitrary units).
(*a*) Nickel B : $\delta T_{\max} \simeq 20°$c, $H = +1{\cdot}40$ Oe, $\delta I = 0{\cdot}38$ e.m.u./mm.
(*b*) Nickel C : $\delta T_{\max} \simeq 1{\cdot}3°$c, $H = +1{\cdot}41$ Oe, $\delta I = 0{\cdot}0197$ e.m.u./mm.
(*c*) Nickel C : $\delta T_{\max} \simeq 1{\cdot}3°$c, $H = +66{\cdot}5$ Oe, $\delta I = 0{\cdot}0282$ e.m.u./mm.

The values of $\delta I/\delta T$ obtained from these results must be corrected for the demagnetizing effect. A change in magnetization at constant applied field H_{a} produces a change in the demagnetizing field and hence a change in the effective field H given by $H = H_{\mathrm{a}} - NI$ where N is the demagnetizing factor for the specimen. The measured change in intensity δI is then $\delta I = (\partial I/\partial T)_H \delta T + (\partial I/\partial H)_T \delta H$ and $\delta H = -N\delta I$, since H_{a} is constant; therefore $\delta I = (\partial I/\partial T)_H \delta T - (\partial I/\partial H)_T N \delta I$ and, if N is small, $(\partial I/\partial T)_H = (\delta I/\delta T)_{H\mathrm{a}}\{1 + N(\delta I/\delta H)_T\}$. When the changes in I are reversible, $(\delta I/\delta H)_T = \kappa_{\mathrm{r}}$, where κ_{r} is the reversible susceptibility and

$$(\partial I/\partial T)_H = (\delta I/\delta T)_{H\mathrm{a}}(1 + N\kappa_{\mathrm{r}}). \qquad \ldots\ldots(9)$$

If the change in magnetization is irreversible then the term $(\delta I/\delta H)_T$ will represent some 'incremental' susceptibility. The correction involved in eqn. (9) is of the order of 6% at $H = +0{\cdot}5$ oersted falling to 1% at $H = \pm 10$ oersted.

In fig. 3, $-T(\partial I/\partial T)_H = (\partial Q'_{\mathrm{rev}}/\partial H)_S$ is plotted against the field, corrected for demagnetizing factor, for $\delta T_{\max} \simeq 20°$ c, 5° c, and 1·3° c; the results for $\delta T_{\max} \simeq 20°$ c and 5° c are given only where the results show a linear relation between change in magnetization and temperature. The results for $\delta T_{\max} \simeq 1{\cdot}3°$ c given in figs. 3 and 4 satisfy the conditions of reversibility given in § 3 over the whole of the field range of ± 120 oersted. As further analyses of the results may be carried out by others, fuller details are given in table 2, where experimental results are given at suitable intervals of field, over the range -50 to $+50$ oersted.

Table 2. Values of $-T(\partial I/\partial T)_H$, I and κ_r, taken from the best Curves drawn through Experimental Points

H	I	κ_r	$-T(\partial I/\partial T)_H$ $\times 10^{-2}$	H	I	κ_r	$-T(\partial I/\partial T)_H$ $\times 10^{-2}$	H	I	κ_r	$-T(\partial I/\partial T)_H$ $\times 10^{-2}$
−50·0	−424·0	0·70	1·05	0	−198	14·0	−0·18	2·5	284·0	7·90	−0·38
−45·0	−420·2	0·72	1·04	0·1	−180	14·7	−0·19	5·0	336·8	4·56	−0·27
−40·0	−416·0	0·75	1·01	0·2	−125	15·3	−0·20	7·5	354·0	3·14	−0·34
−35·0	−411·8	0·80	0·94	0·3	− 42	15·6	$-0{\cdot}21_5$	10·0	366·7	2·38	$-0{\cdot}43_5$
−30·0	−407·3	0·92	$0{\cdot}84_5$	0·4	27	15·6	$-0{\cdot}27_5$	15·0	383·3	1·60	−0·55
−25·0	−402·0	0·98	$0{\cdot}75_5$	0·5	77	15·0	−0·33	20·0	392·5	1·23	−0·65
−20·0	−395·5	1·09	0·65	0·6	112	14·4	$-0{\cdot}37_5$	25·0	399·5	1·02	$-0{\cdot}74_5$
−15·0	−388·0	1·43	0·51	0·8	156	13·1	$-0{\cdot}42_5$	30·0	405·5	0·91	−0·83
−10·0	−375·8	2·15	0·35	1·0	187	12·2	$-0{\cdot}41_5$	35·0	410·9	0·80	−0·87
− 7·5	−367·0	2·65	0·28	1·5	231	10·4	$-0{\cdot}36_5$	40·0	415·7	0·75	−0·93
− 5·0	−355·0	3·35	0·17	2·0	262·5	9·1	−0·40	45·0	420·0	0·72	−0·97
− 2·5	−322·0	5·0	0·01					50·0	424·0	0·70	−1·00

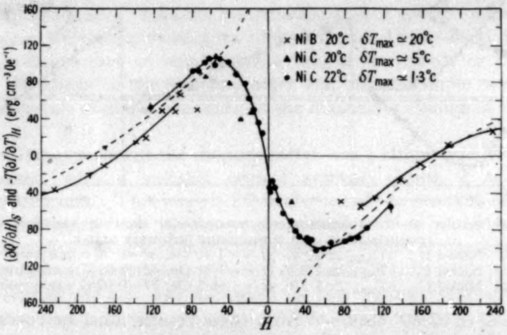


Fig. 3. Derived magneto-caloric effect in annealed nickel (present measurements) and directly observed magneto-caloric effect (Bates and Davis).
Broken line $(\partial Q'/\partial H)_S$ at 18°c (Bates and Davis) ; full lines $-T(\partial I/\partial T)_H$ from magnetometer experiments.

§ 6. DISCUSSION

(i) *Reversible Heat Changes*

Considering only the reversible heating effect, it will be seen that $(\partial Q'_{\mathrm{rev}}/\partial H)_S$ obtained from $-T(\partial I/\partial T)_H$, is in reasonable agreement with the results of Bates and Davis for field values greater than about 50 oersted, and it can be assumed that the magneto-caloric effect in this region is mainly reversible. For fields less than 60 oersted the $(\partial Q'_{\mathrm{rev}}/\partial H)_S$ curve is shown in somewhat more detail in fig. 4. The noticeable 'kink' in the $(\partial Q'_{\mathrm{rev}}/\partial H)_S$ curve is unlikely to be the result of experimental errors as such, as fourteen consistent readings were taken in the region $H=0$ to 5 oersted. It is in this range of field that the correction for demagnetizing factor (eqn. (9)) is as high as 6% and it is possible that the irregularity may be caused by some error in this correction. Some effect of this nature might have been expected near the coercive field ($H=0{\cdot}38$ oersted) where the magnetization changes rapidly, but the irregularity is spread over the range $H=0$ to 5 oersted. It is of course always possible that some small part of the measured $(\partial I/\partial T)_H$ is irreversible but this was certainly not noticeable in the experimental readings.

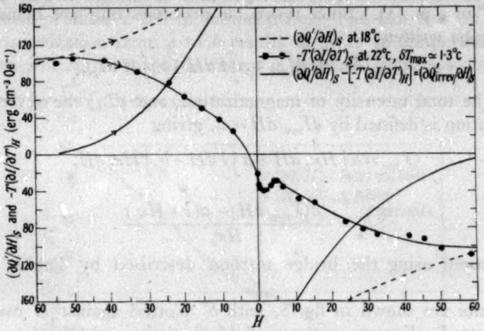


Fig. 4. Derived magneto-caloric effect at low fields.

There are two 'tests' which may be applied to the 'reversible' curve, as a check of its overall correctness. The first is that the net heating or cooling over a cycle, as given by the (Q'_{rev}, H) curve, should be zero; the value obtained from the experimental results (fig. 5) is -150 erg cm^{-3} which is less than 5% of the maximum value of $Q'_{\text{rev}}=3300$ ergs and represents a constant negative

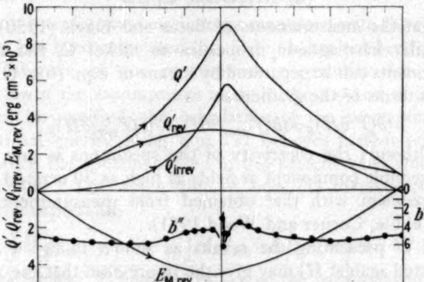


Fig. 5. Magneto-caloric effect in annealed nickel.

Q' = heat developed (Bates and Davis), $Q'_{\text{rev}} = \int_{-50}^{H} -T\,(\partial I/\partial T)_H\,dH$ (Nickel C), $Q'_{\text{irrev}} = Q' - Q'_{\text{rev}}$,

$$(\Delta E_M)_{\text{rev}} = \int_{-50}^{H} H\,dI_{\text{rev}} - Q'_{\text{rev}}.$$

error of about 2% in the results. This is probably experimental error but it may be of some significance that the effect of the irregularity in the low field region of the $((\partial Q'_{\text{rev}}/\partial H)_S, H)$ curve is to produce a net cooling of the same order.

A comparison of the (Q'_{rev}, H) curve with the (Q', H) curve of Bates and Davis indicates that in the field range -50 to $+50$ oersted about one-third of the maximum heat evolved is associated with reversible changes in magnetization.

The second analysis of the $(\partial Q'_{\mathrm{rev}}/\partial H)_S$ results involves the evaluation of the factor b'' of eqn. (3). Since reversible processes only are being considered, eqn. (3) may be written

$$Q'_{\mathrm{rev}} = a\int H\,dI_{\mathrm{rev}} + a\int I\,dH + b''\int H\,dI_{\mathrm{rev}},$$

where I is the total intensity of magnetization, and dI_{rev} the reversible change in magnetization is defined by $dI_{\mathrm{rev}}/dH = \kappa_{\mathrm{r}}$, giving

$$Q'_{\mathrm{rev}} = a\int H\kappa_{\mathrm{r}}\,dH + a\int I\,dH + b''\int H\kappa_{\mathrm{r}}\,dH,$$

and

$$b'' = \frac{(\partial Q'_{\mathrm{rev}}/\partial H) - a(I + H\kappa_{\mathrm{r}})}{H\kappa_{\mathrm{r}}}. \qquad \ldots\ldots(10)$$

κ_{r} was measured using the bridge method described by Tebble and Corner (1950).

The results are shown in fig. 5, with b'' plotted against H over the range $H = \pm 50$ oersted. Except near zero field the values obtained are reasonably constant, with an average value of $b'' = -5{\cdot}1$. This may be compared with the values suggested by Stoner and Rhodes, $b = -4{\cdot}5$ for rotational effects, and $b = -1{\cdot}2$ for reversible boundary movements. From these results, therefore, it would seem that a major part of the magneto-caloric changes in this region is associated with rotational processes.

(ii) *Irreversible Effects*

Assuming that the measurements of Bates and Davis (1950) were made on a specimen similar in magnetic properties to nickel C, the irreversible and reversible components can be separated by means of eqn. (6): $Q' = Q'_{\mathrm{irrev}} + Q'_{\mathrm{rev}}$ as in fig. 5, or in terms of the gradients

$$(\partial Q'/\partial H)_S = (\partial Q'_{\mathrm{irrev}}/\partial H)_S + (\partial Q'_{\mathrm{rev}}/\partial H)_S \qquad \ldots\ldots(11)$$

as in fig. 4. Although the coercivity of the specimens is low, there is still a measurable irreversible component at fields as high as 30 oersted (fig. 5), a result which is in agreement with that obtained from measurements on reversible susceptibility (Tebble, Corner and Wood 1951).

The method of presenting the results as shown in fig. 4 ($(\partial Q'/\partial H)_S$ and $(\partial Q'_{\mathrm{rev}}/\partial H)_S$ plotted against H) may give the impression that the magneto-caloric effect produced by irreversible changes in magnetization near zero field is of a different order from that associated with reversible processes. A clearer indication of the significance of the results is obtained from a comparison of $(\partial Q'_{\mathrm{rev}}/\partial I_{\mathrm{rev}})_S$ and $(\partial Q'/\partial I)_S$ calculated from the present results and from those of Bates and Davis respectively (fig. 6). The uncertainties in these calculations are large but it is obvious that there is no difference in order of magnitude; the corresponding $(\partial Q'_{\mathrm{irrev}}/\partial I_{\mathrm{irrev}}, H)$ curve should thus be of the same form as those in fig. 6, i.e. the heat change associated with a given change in magnetization is of the same order whether the change is reversible or irreversible but, since $\delta Q'_{\mathrm{irrev}} = (\partial Q'_{\mathrm{irrev}}/\partial I_{\mathrm{irrev}})_S \delta I_{\mathrm{irrev}}$, the large changes in magnetization at very low fields produce a correspondingly large change in temperature. In low positive fields the irreversible changes in magnetization are accompanied by a considerable fall in temperature. The impression may sometimes be given that an irreversible process must result in an increase in temperature; this is erroneous and there are several examples of irreversible processes which produce a fall in temperature. The fact that an irreversible change in magnetization from *a* to *b* results in a fall in temperature indicates that the change in internal magnetic energy is greater than the associated changes in energy given by $\int_a^b H\,dI$ (eqn. (2)).

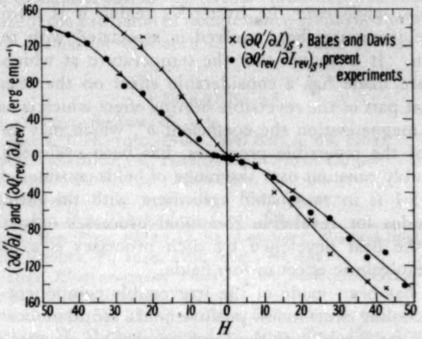


Fig. 6. Variation of $(\partial Q'/\partial I)_S$ and $(\partial Q'_{\rm rev}/\partial I_{\rm rev})_S$ with field.

The change in internal magnetic energy associated with the reversible change in magnetization has been calculated (fig. 5) but it has not been possible to make a satisfactory estimate of the corresponding quantity for the irreversible component. Even if the assumptions made in the method of estimating $Q'_{\rm irrev}$ given above can be justified, the calculation of the corresponding change in irreversible internal energy using eqn. (2) involves further errors due to the differences between the specimen used in the present work (nickel C) and that of Bates and Davis. The application of eqn. (2) involves the evaluation of $\int H\,dI_{\rm irrev}$ from the expression

$$\int H\,dI_{\rm irrev} = \int H\,dI - \int H\,dI_{\rm rev} = \int H\,dI - \int H\kappa_r\,dH, \qquad \ldots\ldots(12)$$

with appropriate limits. Using the values of the magnetization and reversible susceptibility given in table 2, this calculation has been carried out satisfactorily for the specimen nickel C but the results obtained by substituting the values so obtained in eqn. (2) are of little significance, as the method involves the assumption that the same method of estimating $\int H\,dI_{\rm irrev}$ carried out for the specimen of Bates and Davis would give closely similar results. This assumption cannot be justified on the information available and there is no indication of the extent to which the results might be expected to differ.

The problems which have arisen in this separation of the reversible and irreversible change in internal magnetic energy confirm the suggestion made in the introduction to this paper, that the same specimen should be used in the different measurements on the magneto-caloric effect. The importance of this cannot be overemphasized and it is most desirable that a series of co-ordinated investigations should be made on the same specimen; in view of the difficulties caused by demagnetizing effects in cylindrical rods these experiments should be carried out with ellipsoidal specimens.

§7. CONCLUSION

From measurements of the changes in magnetization accompanying reversible temperature changes in annealed nickel a separation of the reversible and irreversible parts of the magneto-caloric effect has been made. It is found that in relatively high fields (greater than 50 oersted) the magneto-caloric effect is mainly reversible, while over the hysteresis region it is estimated that about one-third of the maximum heat evolved is associated with reversible changes in magnetization. It is shown that the temperature at which magneto-caloric measurements are made has a considerable effect on the results. By making allowance for that part of the reversible heating effect which is caused by changes in the intrinsic magnetization the coefficient b'', which may be associated with the remainder of the reversible processes, has been evaluated. The value so obtained is tolerably constant over the range of fields considered and the average value of $b''=-5{\cdot}1$ is in reasonable agreement with the estimate $b=-4{\cdot}5$ of Stoner and Rhodes for reversible rotational processes only; it is suggested, therefore, that the heat developed by such processes is a major part of the reversible magneto-caloric effect in low fields.

An estimate has been made of the irreversible component and it is shown that the marked cooling effect in low positive fields, found in most magneto-caloric measurements, is associated with the large irreversible changes in magnetization in that region.

APPENDIX

A uniformly magnetized ellipsoid, with semi-principal axes a, b, c oriented along $0x$, $0y$ and $0z$ respectively, is uniformly magnetized parallel to the x-axis with magnetic moment M. It can be shown (see for example Stratton 1941, p. 257) that the potential Ω at a point (x, y, z) outside the ellipsoid is given by

$$\Omega=\frac{3Mx}{2}\int_{\xi}^{\infty}\frac{ds}{(a^2+s)R_s} \qquad \ldots\ldots(\text{A } 1)$$

where $R_s=\{(a^2+s)(b^2+s)(c^2+s)\}^{1/2}$, and ξ is the parameter of the family of ellipsoids confocal with the given ellipsoid for which $\xi=0$.

After carrying out the integration in (A 1) for a prolate ellipsoid with $b=c\ll a$, the component of the field at $(x, y, 0)$ parallel to $0x$, to a sufficient approximation, is found to be

$$H_x=-\frac{3M}{a^3}\frac{\partial}{\partial x}\left\{x\left[\ln\left(\frac{(a^2+\xi)^{1/2}+a}{\xi^{1/2}}\right)-\frac{a}{(a^2+\xi)^{1/2}}\right]\right\}, \qquad \ldots\ldots(\text{A } 2)$$

the relations for ξ simplifying to

$$\xi/a^2=\tfrac{1}{2}[n^2+m^2-1+\{(n^2+m^2-1)^2+4m^2\}^{1/2}], \qquad \ldots\ldots(\text{A } 3)$$

where $n=x/a$ and $m=y/a$. (The sign of the square root terms in these expressions is taken as positive.) Writing $\xi/a^2=u^2-1$ and performing the differentiation, (A 2) and (A 3) give

$$-\frac{a^3}{3M}H_x=\tfrac{1}{2}\ln\left(\frac{u+1}{u-1}\right)-\frac{1}{u}-\frac{n^2}{u^3(2u^2-m^2-n^2-1)}. \qquad \ldots\ldots(\text{A } 4)$$

For a point on the x-axis ($m=0$) this becomes

$$-\frac{a^3}{3M}H_x=\tfrac{1}{2}\ln\left(\frac{n+1}{n-1}\right)-\frac{n}{n^2-1}. \qquad \ldots\ldots\text{(A 5)}$$

ACKNOWLEDGMENTS

The authors wish to express their thanks to Professor R. Whiddington, under whose direction this work was commenced, to Professor E. C. Stoner for the generous assistance he has given in the preparation of the paper, to Dr. P. Rhodes for many helpful suggestions, to the Mond Nickel Co. Ltd., who supplied the specimen and to Mr. W. A. Ludbrook who made the ellipsoids. Acknowledgment for the provision of maintenance grants is due to the Department of Scientific and Industrial Research (J. E. W.) and to the London County Council (J. J. F.).

REFERENCES

Bates, L. F., 1949, *J. Phys. Radium*, **10**, 353.
Bates, L. F., and Davis, J. H., 1950, *Proc. Phys. Soc.* A, **63**, 1265.
Bozorth, R. M., and Chapin, D. M., 1942, *J. Appl. Phys.*, **13**, 320.
Bush, H. D., and Tebble, R. S., 1948, *Proc. Phys. Soc.*, **60**, 370.
Hardy, T. C., and Quimby, S. L., 1938, *Phys. Rev.*, **54**, 217.
Okamura, T., 1936, *Sci. Rep. Tohoku Imp. Univ.*, **24**, 745.
Stoner, E. C., 1950, *Rep. Prog. Phys.*, **13**, 83 (London : Physical Society).
Stoner, E. C., and Rhodes, P., 1949, *Phil. Mag.*, **40**, 481.
Stratton, J. A., 1941, *Electromagnetic Theory* (New York : McGraw-Hill) p. 257.
Sucksmith, W., 1945, *J. Sci. Instrum.*, **22**, 129.
Tebble, R. S., and Corner, W. D., 1950, *Proc. Phys. Soc.* B, **63**, 1005.
Tebble, R. S., Corner, W. D., and Wood, J. E., 1951, *Proc. Phys. Soc.* B, **64**, 753.
Tebble, R. S., Skidmore, I. C., and Corner, W. D., 1950, *Proc. Phys. Soc.* A, **63**, 739.

The Effect of Temperature of Deformation on the Electrical Resistivity of Cold-worked Metals and Alloys

By T. BROOM*
Division of Tribophysics, Commonwealth Scientific and Industrial Research Organization,
University of Melbourne

Communicated by W. Boas; MS. received 15*th April* 1952

ABSTRACT. A special apparatus has been used to draw wires of various metals at temperatures between −183°c and 100°c, and to measure their resistances at the temperature of drawing. The change of resistivity with deformation was found to depend on the temperature of deformation relative to the recrystallization temperature. The results explain qualitatively the apparently anomalous resistivity increases previously reported for tungsten and molybdenum. It is suggested that a unified theory of the effect of deformation on the resistivity of both pure metals and alloys can be based on stacking faults.

§1. INTRODUCTION

The increase in electrical resistivity which occurs on cold-working cubic metals and alloys indicates that the deformation generates lattice disturbances which can scatter conduction electrons. The work described in this paper was undertaken in an attempt to acquire further knowledge of these disturbances by the study of changes of resistivity.

---

* Now at Department of Metallurgy, University of Birmingham.

Various workers have determined the changes of resistivity of metals and alloys on deformation at room temperature. A selection of the best available data is given in the Appendix.

One of the most striking features of these data is the very large increase in resistivity reported for tungsten and molybdenum (50 and 18% respectively), compared with the small increase given for copper, aluminium, iron and nickel (2%). The high values for tungsten and molybdenum are not characteristic of the body-centred cubic structure, as iron, which is also body-centred cubic at room temperature, gives a small increase in resistivity. The method of manufacture of tungsten and molybdenum wires (sintering) is probably not significant, as Middleton, Pfeil and Rhodes (1949) found that platinum wires made from a sintered compact gave the same change of resistivity on deformation as those made from an ingot. Experiments by the author showed the same behaviour in copper. Finally, Geiss and van Liempt (1927) showed by measurements of the temperature coefficient of resistance that the large resistivity changes were not due to the formation of internal cracks or fissures. It can therefore be concluded that the large increases in resistivity of tungsten and molybdenum result from disturbances within their crystal lattices.

It was thought that the 'anomalous' behaviour of tungsten and molybdenum might be connected with their high melting points and recrystallization temperatures. Accordingly, experiments have been carried out to see whether the temperature of deformation relative to the recrystallization temperature does indeed determine the magnitude of the resistivity change on deformation. Since the work was commenced, Molenaar and Aarts (1950) have published a note on the change in resistivity of copper, silver and aluminium wires extended at liquid air temperature. The relatively large changes they found support this idea.

It should be pointed out at this stage that the wide variations found in alloys (see Appendix) cannot be explained on this basis since alloys with almost the same melting points and recrystallization temperatures (cf. copper–aluminium and the copper–nickel alloys) give widely different resistivity increases (about 30 and 3% respectively).

Furthermore, it is necessary to examine the basis on which comparisons are to be made between metals. The change in resistivity produced by a given deformation may be expressed as an absolute change $\Delta\rho$ or as a fractional change $\Delta\rho/\rho$ relative to the resistivity ρ of the annealed metal. It is reasonable to assume that in deformed metals the extra resistance due to the scattering of conduction electrons from lattice disturbances is related to the resistance of the undeformed lattice. Accordingly, fractional changes have generally been used to compare different metals. Such comparisons are, of course, only qualitative, as the resistivity of annealed metals varies with temperature, so that fractional changes are dependent on the temperature of measurement.

§ 2. EXPERIMENTAL

Wire-drawing was chosen as the method of deformation so that direct comparison could be made with earlier work. Wire-drawing also has the advantage that much higher deformations can be obtained than are possible by simple extension. All the wires were drawn from an initial diameter of 0·072 in. to a final diameter of 0·022 in., unless repeated breakage made this impossible.

(i) *Materials*

The compositions and annealing treatments of the metals selected for study are listed in table 1.

Aluminium, copper, nickel, iron and 50/50 silver–gold were chosen as being typical of metals which give increases of resistivity of the order of 2% when deformed at room temperature. An α-brass was included because it gives a relatively high increase of resistivity on deformation at room temperature. The ordered alloy Ag$_3$Mg was included in order to study the effect of temperature of deformation on the increase in electrical resistivity of super-lattices.

Table 1. Metals used: Analyses and Annealing Treatments

Metals	Analyses	Heat treatment
Aluminium (spectrographic)	0·002% Mg; <0·001% Si; <0·0005% Fe, Cu.	500°c, 2 h., furnace cooled
Copper	<0·01% Cd, Pb, Ag, Si; $\leqslant$0·01% Mg, Fe.	600°c, 2 h., furnace cooled
Nickel	0·03% C; 0·01% Co; 0·12% Mg; <0·05% Mn, Cr, Cu.	600°c, 2 h., furnace cooled
Iron (spectrographic)	0·002% Cu; 0·005% Ni.	600°c, 2 h., furnace cooled
50/50 Silver–gold	66·13% Au.	600°c, 2 h., furnace cooled
75/25 Brass	74·5% Cu; 0·1% Fe; 0·05% Sn; <0·01% Pb; $\leqslant$0·01% Mn.	600°c, 2 h., furnace cooled
Ag$_3$Mg	6·68% Mg; 93·18% Ag; <0·1% Pb; <0·01% Cu; <0·05% Fe, Mn.	500°c, 2 h., furnace cooled 500°c, 2 h., quenched

(ii) *Apparatus*

To avoid recovery effects it was necessary to measure the resistance of each deformed wire at a temperature equal to or less than the temperature of deformation. Therefore the wires were drawn at a constant temperature and the resistance measurements made without removing the wires from the constant temperature bath. The apparatus used is shown in fig. 1. The wire A was drawn upwards through a die B by means of a chuck C which was raised by a screw mechanism. The double-walled vessel D containing a liquid at the desired operating temperature was attached to the chuck and raised with it as drawing proceeded, thus maintaining the die at the correct temperature.

The baths used were: liquid oxygen (−183°c), solid carbon dioxide and acetone (−78·5°c), ice and water (0°c) and boiling water (100°c). Bath temperatures were checked using a chromel–alumel thermocouple and a millivoltmeter. Low-temperature calibration was carried out using carefully prepared baths of boiling liquid oxygen (−183°c) and of pure acetone mixed with solid carbon dioxide (−78·5°c). Stirrer, thermocouple and immersion heater have been omitted from fig. 1. When operating at −78·5°c and −183°c the double-walled vessel of the drawing apparatus was evacuated to a pressure of less than 0·01 mm Hg. At −78·5°c the sleeves of the cover of the vessel, which normally slid on the columns F supporting the die block E, tended to ice up and stick. This was overcome by winding them with electrical heaters so that they could be warmed intermittently. Curiously, no difficulty of this kind was experienced when operating at −183° c.

The block E holding the die was electrically insulated from its supports F and connected to one current terminal of a Kelvin Double Bridge. The other current terminal was connected to the chuck so that a current could be passed through the wire. Two potential knife edges G connected to the bridge and separated by a known distance (20 cm) were mounted in the apparatus so that they could be moved across on to the wire when a sufficient length had been drawn. The use of a Kelvin Double Bridge avoided any errors due to variable contact resistances at the potential knife edges, and the usual current reversal procedure was followed in order to eliminate the effect of thermal e.m.f's. For simplicity the electrical arrangements are only shown schematically in fig. 1.

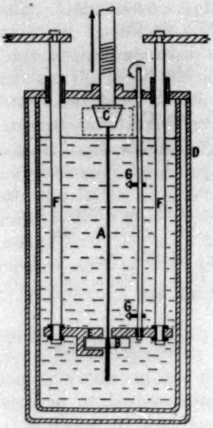


Fig. 1. Apparatus used for drawing wires at controlled temperatures.

In order to obtain large deformations it was necessary to draw the wire successively through several dies. After the resistance of a freshly drawn wire had been measured, the wire was drawn completely through the die and quickly transferred to a long Dewar flask containing liquid at the same temperature as that in the wire-drawing apparatus. The upper end protruded about 5 cm from the top and was swaged in a vertically mounted swaging machine until it would pass through the next die. This die, with a short length of the wire threaded through it, was then transferred to the drawing apparatus, and the processes of drawing and measurement repeated. It should be noted that as the wire extended it was necessary to cut lengths from the top end after each reduction. This ensured that the only part of the wire which had not been held at the desired temperature was repeatedly removed. Further, in no case were resistance measurements made on any length of the wire that had not been held at the correct temperature.

As the wires could not be removed from the constant temperature baths it was impracticable to determine the average cross sections of the lengths of wire on which the resistance measurements were made. The procedure used was to measure the average diameter of the short lengths of wire cut from the end. This was done using a drum micrometer which read to 0·0001 in.

An unsatisfactory feature of the drawing technique was revealed in experiments at −78·5° c and −183° c. It was impossible at these temperatures to draw wires of some metals to any great extent owing to repeated fractures of the wire between the surface of the refrigerant and the chuck. While this was probably due, in part, to abnormally high friction in the die, it should be noted that failure of this kind would be expected if recovery reduced the strength of the exposed upper end of the wire. Because of these failures no results could be obtained at −78·5° c and −183° c for spectrographic aluminium, and only small reductions were possible with some other metals.

Tungsten carbide dies were used throughout the present work and generally no lubricant was employed. However, in a few cases, drawing at −183° c seemed to be facilitated by the use of flake graphite smeared on the die. The reduction in cross-sectional area per die was about 20%.

The resistances of wires drawn at 100° c were measured not only at this temperature but also at 0° c and at −183° c. For this purpose a simple four-point contact device was used with the distance between the potential knife edges adjusted to be the same (20 cm) as that in the drawing apparatus. The wires drawn at 0° c were similarly measured at −183° c, although care had to be taken that all manipulations were carried out below 0° c in order to avoid possible recovery effects.

.B
5
2

§ 3. ERRORS

The major sources of error were as follows:

(i) *Inaccuracies in estimation of average cross section.* When a wire is drawn at room temperature the diameter is often found to vary along the length of the specimen by ± 0·0001 in. Since only short lengths of the wires were measured, errors of this magnitude could be present. The resistivity would then be in error by about 0·03% for wires of 0·072 in. diameter and by about 1% for wires of 0·022 in. diameter. For some of the metals drawn at 100° c and 0° c the fractional change in resistivity $\Delta\rho/\rho$ is small, and such errors in ρ cause appreciable errors in $\Delta\rho$. In these cases a correction was applied by making use of the additional resistance measurement made on these wires at −183° c. The temperature coefficients of resistance α, which do not depend on measurement of wire dimensions, were calculated and Matthiessen's Rule ($\alpha\rho = \text{const.}$) was then used to determine values of $\Delta\rho$.

(ii) *Recovery during transfer from the drawing apparatus to the swaging Dewar.* This will only be important for wires deformed at −78·5° c and −183° c. The transfer time was never greater than two seconds, but it is known that recovery proceeds rapidly even on slight warming. It is difficult to estimate the magnitude of these errors, but their effect will always be to lower $\Delta\rho$.

(iii) *Uncertainties in temperature of drawing.* The dies were always allowed time to attain the desired temperature before drawing was commenced, so that the errors in temperature of drawing are not likely to exceed about 0·5° c for wires drawn at −183° c, 0° c and 100° c. However, it was difficult to maintain the solid CO_2–acetone bath at −78·5° c, and in some cases wires were drawn at temperatures a few degrees higher. It was possible to correct the observed resistances to −78·5° c, but allowance could not be made for the different temperature of drawing.

It will be seen from the form of the curves (figs. 2 to 8) that in this temperature range an increase of 1° c in the temperature of drawing will cause a decrease in $\Delta\rho$ of approximately 1 %.

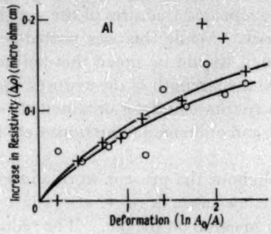


Fig. 2.

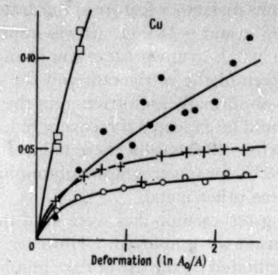


Fig. 3.

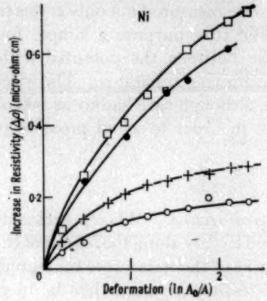


Fig. 4.

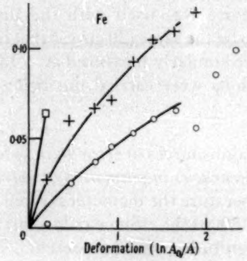


Fig. 5.

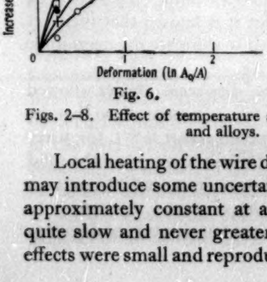


Fig. 6.

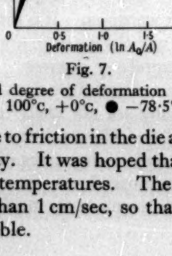


Fig. 7.

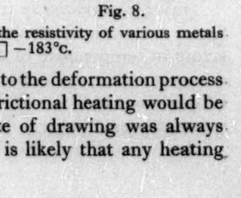


Fig. 8.

Figs. 2–8. Effect of temperature and degree of deformation on the resistivity of various metals and alloys. ○ 100°c, +0°c, ● −78·5°c, □ −183°c.

Local heating of the wire due to friction in the die and to the deformation process may introduce some uncertainty. It was hoped that frictional heating would be approximately constant at all temperatures. The rate of drawing was always quite slow and never greater than 1 cm/sec, so that it is likely that any heating effects were small and reproducible.

§ 4. RESULTS

The variations of $\Delta\rho$ with degree and temperature of deformation are shown in figs. 2 to 8, and the resistivities of the annealed metals at the temperatures considered are given in table 2. Table 2 also shows values of $\Delta\rho/\rho$ at -183° c for the maximum deformation that was reached in each case.

Table 2. Resistivities of Annealed Metals and Fractional Increases on Deformation at -183° c.

Metals	Resistivities ($\mu\Omega$ cm)				Deformation (%)	$\frac{\Delta\rho}{\rho}\times 10^2$
	-183°c	$-78{\cdot}5^\circ$c	0°c	100°c		
Aluminium	0·352	1·55	2·44	3·59	—	—
Copper	0·291	1·06	1·59	2·23	37	38
Nickel	1·77	4·59	7·37	11·56	90	40
Iron	1·09	5·78	9·06	14·73	20	6
75/25 Brass	4·09	5·10	5·84	6·70	90	69
50/50 Silver–gold	9·38	10·10	10·78	11·55	37	4
Ag_3Mg (ordered)	2·23	3·64	4·64	5·84	37	112
Ag_3Mg (disordered)	7·53	8·60	9·25	10·01	50	11

The correction discussed in § 3 (i) has been applied to the ($\Delta\rho$, deformation) curves for copper, aluminium, nickel and iron, deformed at 100° c and 0° c. In these cases the curves of (α, deformation) were quite smooth, and the values of $\rho(0^\circ\,\mathrm{c})-\rho(-183^\circ\,\mathrm{c})$ were found to be independent of the degree of deformation, thus justifying the use of Matthiessen's rule.

The curves for 50/50 silver–gold have been omitted as the changes in resistivity were so small, even at -183° c, that the errors in measurement of wire diameter caused very large scatter. The corrections for measurements at 0° c and 100° c could not be applied as the temperature coefficient is very small and does not vary greatly with deformation.

§ 5. DISCUSSION

The results show that for the metals investigated the increase in electrical resistivity due to cold work depends markedly on the temperature of deformation. At any one temperature the increase with strain is at first rapid and then becomes more gradual. The effect of decreasing temperature of deformation is to increase the rate of change of resistivity at all strains and to prolong the initial period of rapid increase. It is clear from these results that to compare the behaviour of different metals we must consider the two families of ($\Delta\rho$, strain) curves and not only the resistivity increases at arbitrarily chosen temperatures and strain values. However, even without a detailed analysis, it is clear from the values of $\Delta\rho/\rho$ (table 2) for a temperature (-183° c) far below the recrystallization temperatures of the metals concerned that the behaviour of tungsten and molybdenum at a correspondingly low temperature (room temperature) is not anomalous.

Since it must be possible to interpret the present data in terms of the mechanism of plastic deformation, these data are, in principle, capable of yielding information concerning the nature of the lattice disturbances produced by cold work. The most obvious approach to this problem is to find the extent to which resistivity data can be explained in terms of those lattice disturbances which are already known to exist in deformed crystals. Of these, dislocations will be discussed first.

The similarity in shape of the ($\Delta\rho$, strain) and the strain-hardening curves and the effect of temperature on these curves give qualitative support for the idea that dislocations are responsible both for resistivity change and strain hardening. However, a detailed explanation of how dislocations can play this dual role is by no means straightforward. Previous work (see, for instance, Barrett 1943, Molenaar and Aarts 1950) has shown quite clearly that appreciable recovery of resistivity takes place at much lower temperatures than recovery of strain hardening. Thus, if dislocations are wholly responsible for the increase in resistivity then all dislocations cannot contribute equally to the resistivity and to hardness. The only calculations of the resistivity increase due to dislocations (Mackenzie and Sondheimer 1950, Landauer 1951) are based on the assumption of equal contribution of all dislocations, and hence these analyses are not fully satisfactory in spite of the approximate agreement obtained with experiments on copper.

The rapid recovery at low temperatures led Molenaar and Aarts to the conclusion that different mechanisms were responsible for the increase in resistivity and for strain hardening. Mott (1951) has suggested that vacant lattice sites rather than dislocations are mainly responsible for the resistivity increase on cold work. He has pointed out that vacancies would be effective in increasing resistivity and would be very mobile at low temperatures, thus allowing recovery to occur. This has also been discussed by Blewitt, Taylor and Coltmann (1951), Druyvesteyn and Manintveld (1951) and Broom (1952). Seitz (1950) has shown how vacancies can be generated by various processes involving the movement of dislocations. These ideas can obviously be extended to account for the large increases in resistivity on deformation at low temperatures.

While it is agreed that a major part of the increase in resistivity may be due to lattice vacancies, a more careful analysis of the dislocation hypothesis suggests another possibility of accounting for the recovery data.

Dislocations could be responsible for the increases both in resistivity and in hardness if the various features of the dislocation arrays influenced these two properties differently. Indeed, it would be surprising if all the features (number, shapes, sizes, arrangement, etc.) of the dislocation loops were to influence the two properties in the same way. A particular feature of dislocations in face-centred cubic lattices is their dissociation into two half-dislocations with the formation of a fault. As Heidenreich and Shockley (1948) have pointed out, the extent of the region of faulty stacking between the half-dislocations is dependent on their mutual repulsion and on an opposing force tending to reduce the volume of the fault, as this has a higher configurational energy than the rest of the lattice. If the energy difference between the ABAB (hexagonal close-packed) and the ABCABC (cubic close-packed) packing sequence is small, the fault associated with the two half-dislocations will be large, and vice versa. Heidenreich and Shockley have shown that faults contribute to strain hardening by interfering with the passage of dislocations along slip planes which intersect the fault. Similarly, since a fault represents a disturbance in the periodicity of the lattice it hinders the passage of conduction electrons. Provided the faults are sufficiently widely separated, their effect on resistivity will simply be additive, i.e. it is the total area of the fault projected on the plane normal to the direction of the current that controls the resistivity increase. However, their effect on yield stress would probably not be additive. To a first approximation (neglecting interaction between faults) a series of faults merely represents a succession of obstacles to the passing dislocation, and an applied stress sufficient to move the dislocation through one fault suffices for the whole series. However, the half-dislocations make an independent contribution to strain hardening because of their stress fields.

These ideas lead to a second interpretation of the effect of temperature of deformation. The larger increases in resistivity at low temperatures can be attributed to an increase in the size of the faults produced. This conforms with the general observation that the hexagonal close-packed structure becomes increasingly stable with respect to the cubic close-packed structure as the temperature is lowered.

Similarly, the recovery of resistivity at low temperatures can be attributed to an adjustment of the separation of the half-dislocations, thus reducing the size of the fault and its capacity for scattering electrons. This adjustment might be expected to occur at much lower temperatures than those needed to permit dislocations to migrate right out of the lattice. This change in the area of the fault would not be expected to influence the hardness to anything like the extent to which resistivity would be affected. Although the individual faults would no longer intersect as many slip planes, the number of faults would ensure that all slip planes still intersect faults. The contribution to the hardness from the half-dislocations themselves would remain unchanged.

However, the most attractive feature of this fault hypothesis is that it promises to provide a unified picture of the effects of deformation on the resistivity both of pure metals and alloys. It is difficult at present to assess the faulting tendencies of different metals and alloys. One criterion is the hardening of latent slip planes, which has been attributed to the formation of faults (Heidenreich and Shockley 1948). It is interesting to note that the α-solid solutions of copper–aluminium and copper–zinc which show marked unequal hardening of the operative and latent slip planes also give large increases in resistivity on deformation. On the other hand, solid solutions such as 50/50 silver–gold and copper–nickel, where operative and latent slip planes harden equally, give resistivity increases of a few per cent only. X-ray evidence obtained by Barrett (1949) and by Barrett and Barrett (1951) provides further opportunity for correlating resistivity and fault density. They have rated the following alloys in groups of increasing faulting tendency on deformation at room temperature: (i) Al, Cu, Ag; (ii) Cu–8·5% Zn, Cu–1·89% Si; (iii) Cu–30% Zn, Ag–7% Sb; (iv) Cu–4·2% Si, Ag–11% Sn. There is a qualitative correspondence between these results and the values of increase in resistivity given in the Appendix for similar alloys.

Finally, a few experiments have been carried out on cobalt, which is known to contain stacking faults in the annealed state (Edwards and Lipson 1942), and on cobalt–nickel alloys. It would be expected that the number of faults produced by deformation would be a maximum near the boundary between the hexagonal close-packed and the cubic close-packed structures (~30% Ni). The alloys were deformed 20% by drawing at room temperature. The increases in resistivity are shown in table 3. These results confirm the prediction that the maximum increase in resistivity occurs in the 30% nickel alloy.

Suggestions have been made that destruction of short-range order is responsible for the large increases in resistivity given by some alloys on deformation. However, in the present experiments it has been shown that 50/50 silver–gold gives a very

Table 3. Resistivity Changes of Cobalt–Nickel Alloys after 20% Deformation by Wire Drawing at Room Temperature.

Composition (% Ni)	0	30	40	100
$\Delta\rho/\rho$ (%)	4	7	2	1

small increase in resistivity despite the existence of short-range order as shown by Guinier (1945). Further, the experiments of Jones and Sykes (1938) on $AuCu_3$ indicate that short-range order (or long-range order with small domain size) has little effect on resistivity.

In conclusion, it is felt that an understanding of the effect of deformation on the electrical resistivity of pure metals and of alloys is dependent upon an assessment of the relative importance of three types of lattice disturbance—dislocations, vacancies and stacking faults. The need for a quantitative treatment of the effect of stacking faults on resistivity is apparent.

ACKNOWLEDGMENTS

The author is indebted to the Goldsmiths' Company for a Dominion Travelling Scholarship during the tenure of which much of the work was carried out. Thanks are also due to Dr. W. Boas for his advice and constant encouragement, to Mr. J. S. Bowles for his assistance in preparation of the manuscript for publication, and to Mr. E. D. Hondros for assistance in the experimental work.

REFERENCES

Barrett, C. S., 1943, *Structure of Metals* (New York: McGraw-Hill), p. 375; 1949, *Cold Working of Metals* (Cleveland: Amer. Soc. Metals), pp. 65–98.
Barrett, C. S., and Barrett, M. A., 1951, *Phys. Rev.*, **81,** 311.
Blewitt, T. H., Taylor, W., and Coltmann, R. R., 1951, *Phys. Rev.*, **82,** 769.
Broom, T., 1952, *Austral. J. Sci. Res.* A, **5,** 128.
Clarebrough L. M., and Nicholas, J. F., 1950, *Austral. J. Sci. Res.* A, **3,** 284.
Crampton, D. K., Burghoff, H. L., and Stacy, J. T., 1941, *Trans. Amer. Inst. Min. Met. Engrs.*, **143,** 228.
Dahl, O., 1936, *Z. Met.*, **28,** 133.
Druyvesteyn, M. J., and Manintveld, J. A., 1951, *Nature, Lond.*, **168,** 868.
Edwards, O. S., and Lipson, H., 1942, *Proc. Roy. Soc.* A, **180,** 268.
Geiss, W., and van Liempt, 1927, J. A. M., *Z. Phys.*, **41,** 867.
Guillet, L., and Ballay, M., 1923, *Rev. Met.*, **20,** 398.
Guinier, A. J., 1945, *Proc. Phys. Soc.*, **57,** 310.
Heidenreich, R. D., and Shockley, W., 1948, *Strength of Solids* (London: Physical Society), p. 57.
Jones, F. W., and Sykes, C., 1938, *Proc. Roy. Soc.* A, **166,** 376.
Landauer, R., 1951, *Phys. Rev.*, **82,** 520.
Mackenzie, J. K., and Sondheimer, E. H., 1950, *Phys. Rev.*, **77,** 264.
Middleton, A. B., Pfeil, L. B., and Rhodes, E. C., 1949, *J. Inst. Met.*, **75,** 603.
Molenaar, J., and Aarts, W. H., 1950, *Nature, Lond.*, **166,** 690.
Mott, N. F., 1951, *Proc. Phys. Soc.* B, **64,** 729.
Seitz, F., 1950, *Phys. Rev.*, **79,** 723, 890 and 1002.
Smart J. S., Smith, A. A., and Phillips, A. J., 1941, *Trans. Amer. Inst. Min. Met. Engrs.*, **143,** 272.
Tammann, G., and Dreyer, K. L., 1933 a, *Ann. Phys., Lpz.*, **16,** 111; 1933 b, *Ibid.*, **16,** 657.

APPENDIX

Change of Electrical Resistivity with Deformation at Room Temperature

Cubic metals	Method of deformation	Reduction of cross section (%)	Increase of resistivity (%)	Reference
Aluminium	Drawn	88	2·2	Present work
Copper	,,	87	2·4	Smart, Smith & Phillips 1941
Gold	Rolled	98	1·6	Tammann & Dreyer 1933 a
Iron	Drawn	85	1·3	Present work
Molybdenum	,,	>99	18	Geiss & van Liempt 1927
Nickel	,,	90	3·9	Present work
Palladium	,,	96	4	Tammann & Dreyer 1933 a
Platinum	,,	96	1	,, ,,
Silver	,,	98	5	,, ,,
Tungsten	,,	>99	50	Geiss & van Liempt 1927
Alloys				
Copper 7·5% Al	Drawn	75	27	Crampton, Burghoff & Stacy 1941
10% Au	Rolled	97	1·7	Tammann & Dreyer 1933 b
20% Mn	,,	97	9	,, ,,
20% Ni	Drawn	84	3	Crampton, Burghoff & Stacy 1941
3% Si	,,	84	9	,, ,,
2·4% Sn	,,	84	2	,, ,,
30% Zn	,,	84	21	,, ,,
Gold 15% Ag	Rolled	97	1·4	Tammann & Dreyer 1933 b
50% Ag	Drawn	90	1·1	Present work
Platinum 25% Ir	,,	50	0·5	Guillet & Ballay 1923
Silver 25% Au	Rolled	97	0·8	Tammann & Dreyer 1933 b
30% Cd	,,	97	28	,, ,,
23% Zn	,,	97	19	,, ,,
Alloys Order-disorder type				
Ag_3Mg ordered	Drawn	90	100	Clarebrough & Nicholas 1950
disordered	,,	90	8	,, ,,
Cu_3Au ordered	,,	70	45	Dahl 1936
disordered	,,	70	2	,,
Ni_3Fe ordered	,,	95	35	,,
disordered	,,	95	15	,,
Ni_3Mn ordered	Rolled	95	37	,,
disordered	,,	95	4	,,

Yield Points in Bending Experiments on Zinc Crystals

By M. J. DUMBLETON and B. W. HOWLETT
Metallurgy Department, University of Birmingham

Communicated by A. H. Cottrell; *MS. received* 16*th May* 1952

ABSTRACT. It is shown that sharp yielding can be observed in previously unstrained zinc crystals by means of bending experiments devised to avoid premature overstraining of the specimen. This resolves a difficulty in the interpretation of previous tensile experiments on crystals which show the yield phenomenon.

§ 1. INTRODUCTION

Although it has been shown that sharp yield points can be produced in tensile tests on single crystals of zinc containing nitrogen (Wain and Cottrell 1950), some doubt remains about whether strain ageing is a prerequisite for their observation (Orowan 1949). The absence of yield points in initial tests may mean that the yield phenomenon does not exist in an unworked crystal, a fact which would be of considerable importance to the theory of the yield point, or merely that the yield point is not revealed during the first test on account of stress concentrations. A stress concentration could cause localized yielding before the mean stress in the specimen reached the lower yield stress, in which case a smooth stress–strain curve could result. Plastic straightening of the specimen, which is often produced by the straining and ageing treatment given to it, largely smooths away these stress concentrations, and this could be the reason why strain-aged specimens show yield points.

The purpose of the work described below was to investigate this question. Most previous experiments have been made on tensile specimens, but here difficulties arise from stress concentrations at the grips and from non-axiality of loading. Furthermore, it is difficult to ensure that very soft crystals, such as those of zinc, are not bent or otherwise damaged during the handling necessary to mount them in the testing machine. To overcome these difficulties a machine was designed which allowed single crystal wires of zinc to be tested by bending. In these tests no grips were used, the question of axiality did not arise, and it was not necessary to handle the specimen. This made it possible to prove that the yield phenomenon exists in previously unstrained crystals.

§ 2. APPARATUS

The testing machine is represented diagrammatically in (*a*) and (*b*) of fig. 1. A graduated beam B is supported by, and can pivot about, two needle points which engage in a groove in a steel rod A fixed perpendicularly to the beam axis. The counterweights C and C′ are adjusted to balance the beam when the rider R is at the far end of the scale. The load is transmitted to the specimen S through the knife-edge K fixed on the near end of the beam, while at the far end the point P makes electrical contact with the anvil M of a micrometer.

The specimen projects from a silica quill Q held firmly in a fixed clamp.

§ 3. METHOD

Single crystals of zinc were grown by the Andrade and Roscoe fusion method (1937), in the form of wires of 1 mm diameter. Two grades of zinc were used, 'pure' and 'commercial', the spectrographic analyses of which are given by Wain and Cottrell (1950). In tensile tests, crystals grown from the commercial zinc yielded sharply after strain ageing, but those grown from the pure zinc gave no sharp yields even after several strain-ageing treatments.

A specimen was mounted in the machine, without removing it from the quill in which it was grown, by clamping this quill so that its end was 5 cm away from the knife-edge. The specimen was then pushed gently along the quill with a piece of wire until it projected by a little more than 5 cm; this was the only handling to which it was subjected between growing and testing. Since the total length of the specimen was 8 cm it tilted up until its inner end rested against the top of the inside of the quill.

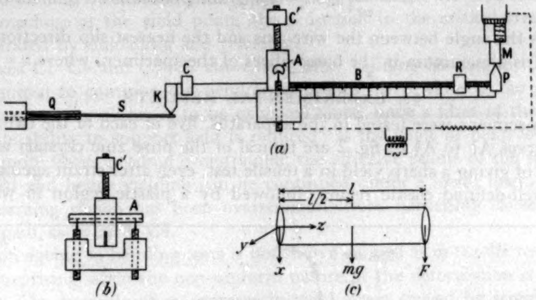


Fig. 1.

To test the specimen the knife-edge was first raised by depressing P with the micrometer, and the rider moved along the graduated beam towards the fulcrum. The knife-edge was then lowered again on to the specimen, thus imposing a known force on it. The corresponding displacement of the end of the specimen was measured by the micrometer, contact between this and the point P being indicated by an electrical circuit. The machine was of the 'dead-loading' type so that, whenever the yield phenomenon occurred, the specimen continued to bend under the constant applied load until the micrometer, which served as a stop, became engaged once more.

After a test had been made, the specimen, its quill, and its clamp could be removed as a unit from the machine and set aside to allow the specimen to recover. At the conclusion of the experiment the orientation of the specimen was determined using an optical goniometer.

§ 4. CALCULATION OF THE RESOLVED SHEAR STRESS

We assume that, for the purpose of calculating the stresses at the place where the crystal emerges from its quill, the weight of the projecting part of it acts at the centre of this part. This reduces the problem to that of a cylindrical cantilever of radius r acted on at distances $l/2$ and l from its point of support by two transverse forces, the weight mg of the projecting part and the force F

.B
5
2

applied through the knife-edge (fig. 1(*c*)). The solution of this problem is well known (e.g. Timoshenko 1934). Three components of stress do not vanish; the shear stress τ_{yz}, the shear stress τ_{xz} which balances the forces on the specimen, and the tensile stress σ_z which balances the couple due to those forces.

It has been shown by Roscoe (1936), who worked with cadmium crystals, that in a bending test in which the resolved shear stress over the slip plane is not constant, it is the maximum rather than the mean value of the shear stress which determines the onset of slip. It turns out that, for the usual experimental conditions, e.g. $l = 5$ cm, $r = 0{\cdot}5$ mm, and $\chi < 60°$, where χ is the angle which the slip plane makes with the wire axis, the maximum shear stresses in the slip direction due to τ_{xz} and τ_{yz} are negligible compared with the maximum shear stress due to σ_z. Hence we may write the resolved shear stress σ as

$$\begin{aligned}\sigma &= \sigma_z \sin\chi \cos\lambda \\ &= -\frac{4}{\pi r^4}\, x(Fl + mgl/2) \sin\chi \cos\lambda,\end{aligned}$$

where λ is the angle between the wire axis and the nearest slip direction. This, of course, is a maximum in the outer fibres of the specimen, where $x = \pm r$.

§ 5. EXPERIMENTAL RESULTS

Ten specimens were tested in the apparatus, five of each of the two kinds of zinc. Curves A1 to A4 in fig. 2 are typical of the pure zinc crystals which are incapable of giving a sharp yield in a tensile test, even after strain ageing. They show a well-defined elastic region, followed by a plastic region in which the

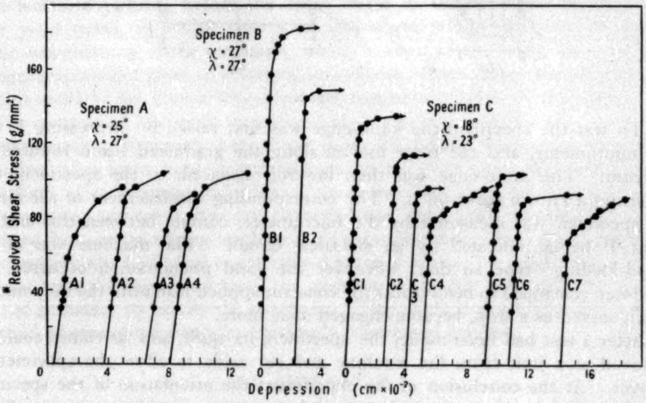


Fig. 2.

* The scale of stress shown gives the resolved shear stress calculated for the outer fibres of a specimen deforming purely elastically.

Curves for Specimen A.

A 1. Initial test on crystal of pure Zn.
A 2. Immediate retest.
A 3. Test taken 24 hours later.
A 4. Immediate retest.

Curves for Specimen B.

B 1. Initial test on crystal of commercial Zn.
B 2. Immediate retest.

Curves for Specimen C.

C 1. Initial test on crystal of commercial Zn.
C 2. Immediate retest.
C 3. Immediate retest.
C 4. Immediate retest.
C 5. Test taken after 44 hours.
C 6. Immediate retest.
C 7. Immediate retest.

rate of work hardening decreases steadily. The raising of the elastic limit by straining is illustrated by curve A2, which was obtained immediately after curve A1, while recovery of the elastic limit is shown in curve A3, obtained 24 hours later. There are no signs in these curves of either strain softening or strain ageing, which are the identifying features of the yield phenomenon.

The curves shown for specimens B and C are typical of crystals of commercial zinc which yield sharply when tested in tension after strain ageing. In the first test on specimen B, shown in curve B1, the specimen started to flow very rapidly under the applied load, the depression of its end being limited only by the anvil of the micrometer. This is represented by an arrow on the curve. On unloading and immediately retesting, flow began at a considerably lower load (curve B2). The fact that the specimen flowed rapidly under an applied load for which no flow could be observed in the first test demonstrates the existence of the yield phenomenon in the material, for this is its most characteristic feature. Although in a dead-loading machine there can be no drop in load of the kind observed in a 'hard' machine at the yield point, the reduction in the critical stress can be demonstrated by unloading and retesting.

Curves C1–C4 and C5–C7 show how, after yielding has begun, the applied load required to continue the bending decreases for a while as the amount of bending increases. This is to be expected since, once a fibre of the specimen yields, its ability to support load is reduced; as the angle of bend is increased, so that more fibres become overstrained, the effective radius of the specimen is decreased and hence yielding can be continued under a lower load. When all the deforming region has been overstrained, work hardening causes the load to rise again, as in curve C4.

Strain ageing in bending tests is not shown so well as in tensile tests. This is not surprising when the non-uniform nature of the deformation is taken into account. An example of an increase in yield stress caused by strain ageing is shown in curve C5, obtained 44 hours after curve C4.

§ 6. CONCLUSIONS

Zinc crystals subjected to bending tests give stress–strain curves similar to those obtained in tensile tests using a dead-loading machine. In susceptible material yield points and strain ageing can be obtained.

Yield points can be observed in unstrained crystals by means of the bend test when this is arranged so that premature overstraining of the specimen due to stress concentrations from the grips, from non-axial loading, and from handling, are avoided.

It is concluded that previous plastic deformation is not essential to the appearance of a sharp yield in zinc crystals, and that failure to observe initial sharp yielding in tensile tests is due to limitations of the experimental technique.

ACKNOWLEDGMENTS

The authors wish to thank Professor A. H. Cottrell for his interest and advice, and Professor D. Hanson for his general supervision and support. The work forms part of a programme supported by the Department of Scientific and Industrial Research through the National Physical Laboratory, to whom grateful acknowledgment is made.

REFERENCES

Andrade, E. N. da C., and Roscoe, R., 1937, *Proc. Phys. Soc.*, **49**, 152.
Orowan, E., 1949, *Trans. Amer. Inst. Min. Met. Engrs.*, **185**, 878.
Roscoe, R., 1936, *Phil. Mag.*, **21**, 339.
Timoshenko, S., 1934, *Theory of Elasticity* (New York : McGraw-Hill), p. 289.
Wain, H. L., and Cottrell, A. H., 1950, *Proc. Phys. Soc.* B, **63**, 339.

Some Observations on the Yield Point in Zinc

By H. L. WAIN

Aeronautical Research Laboratories, Department of Supply, Melbourne, Australia

Communicated by A. H. Cottrell; *MS. received* 30*th June* 1952

ABSTRACT. Evidence is given confirming that the yield point in zinc is due to nitrogen in solution. The theoretical prediction that a yield point should appear in a susceptible crystal without the necessity for any prior deformation is verified experimentally, and an explanation is given of the observation that a yield point may appear in a specimen even though free dislocations have been introduced subsequent to ageing.

Under certain conditions the yield point values in zinc decrease as identical strain-ageing treatments are repeated, and a smooth stress–strain curve is ultimately obtained. This is attributed to an accumulation of work hardening which causes the dislocation density to increase to such an extent that the dislocations can no longer be effectively anchored by the available nitrogen atoms. The difference between this behaviour and that of iron on repeated strain ageing is discussed.

§ 1. INTRODUCTION

CONSIDERABLE attention has been given to the question of the yield point in metals in recent years. A theoretical explanation has been developed (Cottrell 1948, Nabarro 1948, Cottrell and Bilby 1949, Cottrell 1950) which ascribes the effect to the preferential collection of specific foreign atoms in the lattice at dislocations. Once such segregation has occurred the dislocations are rendered immobile until a sufficient stress is applied to break the bonds which are formed between the dislocations and the segregate of foreign atoms, whereupon the freed dislocations move through the lattice at a lower stress level and a yield point is shown. No yield point appears on testing immediately after overstraining; but suitable annealing allows the solute atoms to move again to the dislocations and anchor them (strain ageing). This theory explains the well-known effects in polycrystalline mild steel and has received further confirmation from the fact that yield points have been observed in iron single crystals containing carbon (Schwartzbart and Low 1949, Churchman and Cottrell 1951) and in single crystals of cadmium (Smith 1947, Cottrell and Gibbons 1948) and zinc (Orowan 1934, 1940, Wain and Cottrell 1950).

An investigation into some yield point effects in zinc crystals is described below and includes work on the dissolution of nitrogen in zinc, on the development and magnitude of the yield point, and on the influence of work hardening on the yield point characteristics.

§ 2. METHOD

The material used was spectrographically pure zinc treated to produce nitrogen absorption by melting in an open combustion boat in air (Wain and Cottrell 1950) or under a stream of nitrogen in an electric resistance furnace at constant temperature. In both cases dissolution of nitrogen was facilitated by keeping the molten zinc free from surface films by frequent sprinkling with ammonium chloride flux. After this treatment the zinc was re-melted and cast into $\frac{3}{8}$ in. diameter ingots which were extruded into 1 mm diameter wire. Both the re-melting and casting operations were carried out under a nitrogen atmosphere. Spectrographic analysis revealed that the only metallic impurities in the extruded wires were lead, iron, copper and cadmium, and that the total impurity content was less than 0·003%.

Single crystals were grown from the extruded wire by Andrade and Roscoe's (1937) fusion method, and mechanical tests were made on a modified Polanyi (1925) machine. Such machines are of the rigid frame type and the principle of operation is illustrated in fig. 1. A copper loop was welded on to each end of the single crystal specimen and each loop was engaged by a hook on the testing machine. The upper hook was attached to a beam and the lower to a micrometer screw, and beam and screw were connected through a rigid frame. Incremental loading was applied by turning the micrometer screw; the load on the specimen was determined to within 1 g by an optical measurement of the deflection of the beam, and the elongation of the specimen was given by the movement of the screw measured to 10^{-4} in. Shear stress–strain curves could be calculated from these results provided the load–deflection calibration of the beam and the orientation of the crystal relative to the wire axis were known. The latter property was determined for each crystal by an examination of etch-pit reflections with an optical goniometer (Barrett and Levenson 1940). Particular care was necessary to ensure axiality of loading, and this was achieved by careful alignment of the machine and by interposing a fine chain between the beam and the upper hook engaging the specimen.

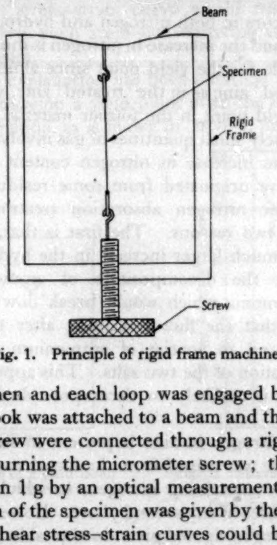


Fig. 1. Principle of rigid frame machine.

The specimen was annealed between tests in a paraffin oil bath at temperatures up to 200 °c, and the design of the machine was such that this could be done without removing the specimen from the machine. The annealing bath consisted of a Dewar flask fitted with an internal heating coil by means of which paraffin oil could be heated and maintained indefinitely at a given temperature to within ±1° c.

§ 3. NITROGEN IN ZINC

Evidence presented previously (Wain and Cottrell 1950) has indicated strongly that the presence of nitrogen is necessary before a yield point can be developed in zinc. This view has been confirmed by gas analysis, using the vacuum fusion technique, of 'as received zinc' and of zinc which had been treated as described in § 2 to produce the yield point effect. The results of these tests are given in the table.

Gas Analysis of Zinc

Sample	Oxygen (% by wt.)	Nitrogen (% by wt.)	Hydrogen (% by wt.)
'As received' zinc	nil	0·0002	0·00037
'Treated' zinc	0·0001	0·0022	0·00048

Thus an increase occurs in both nitrogen and hydrogen when zinc is treated to produce a yield point and the increase in nitrogen is much the larger. Hydrogen can hardly be responsible for the yield point since almost as much hydrogen is present in the 'as received' zinc as in the 'treated' zinc, yet it has not been found possible to develop a yield point in the former material.

In view of the extremely small quantities of gas involved in the above analysis, it is conceivable that the increase in nitrogen content reported above for the 'treated' zinc could have originated from some residual ammonium chloride in the sample after the nitrogen absorption treatment. However this is considered unlikely for two reasons. The first is that, if it had been so, one would have expected a much larger increase in the hydrogen content than was found to occur, since the decomposition of ammonium chloride would presumably liberate ammonia which would break down as: $2NH_3 \rightarrow N_2 + 3H_2$. The second reason is that the flux remaining after the nitrogen absorption treatment has been found to consist of ammonium chloride + zinc chloride —apparently a solid solution of the two salts. This appears to be stable to quite high temperatures and is unlikely to decompose during the vacuum fusion analysis.

An indication that nitrogen in zinc is present in solid solution rather than in the form of zinc nitride was gained from the fact that microscopic examination of single crystals of zinc containing nitrogen, in which yield points had been developed, showed no traces of any second phase. However, it should be pointed out that, even if the nitrogen figures given in the table represent the amount in solid solution, it is unlikely that this percentage would be present in single crystals grown from this material. The reason for this is that the zinc is twice melted subsequent to the nitrogen absorption treatment—once to prepare billets for extrusion and once during the growth of the single crystals—and it would be expected that some nitrogen would be lost during each of these re-melting operations, although both were carried out under nitrogen. Thus the value of 0·0022 wt. % N_2 (0·014 atomic %) must be considered as an upper limit as far as the nitrogen content of the single crystals is concerned. At first sight it is surprising that such minute quantities of a solute atom can produce the yield point effect, but a calculation based on a simple model (Cottrell 1949) shows that only 0·00001 atomic % of solute is required to lock all the dislocations in an annealed crystal (10^8 dislocations/cm^2) and this figure rises to 0·1 atomic % in a severely cold worked metal (10^{12} dislocations/cm^2).

§4. INITIAL DEVELOPMENT OF THE YIELD POINT

Experience gained during the testing of zinc crystals has shown that a yield point is not normally shown in a susceptible specimen until several strain-ageing treatments have been given. There appears to be no theoretical reason for this behaviour since there is believed to be a considerable number of dislocations present in a fully annealed crystal and these should be anchored in the usual way on the first annealing treatment. This phenomenon is considered rather to result from the conditions of testing (Wain and Cottrell 1950) in that the specimen is likely to be somewhat bent during mounting and placing on the testing machine, the load may not be applied axially, and there may be stress concentrations at the specimen grips. These factors may act as stress-concentrators causing yielding to occur at a mean stress less than the lower yield point so that a smooth stress–strain curve would result. Repetition of the strain-ageing cycle removes such stress-concentrating effects by plastic flow so that the yield point appears.

Confirmation of this explanation has been obtained from the fact that it has been found possible to develop a yield point after the first annealing treatment without any prior deformation, as is shown in curve 1, fig. 2. Particular care had been taken in handling this specimen and it had been electropolished before testing in order to remove surface irregularities which may have acted as stress-raisers. Even with this specimen, however, it was found that the yield point values increased on subsequent annealing (curve 3), indicating that some stress concentrations were present when the specimen was first tested in spite of the precautions which had been taken. This production of the yield point on the first annealing treatment is in line with the results on other materials. Cottrell and Gibbons (1948) have reported that yield points sometimes appear after the first annealing treatment in cadmium crystals, whilst Churchman and Cottrell (1951) have found the same effect in single crystals of iron.

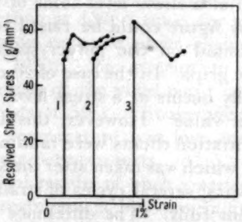


Fig. 2. Development of a yield point without prior plastic deformation.

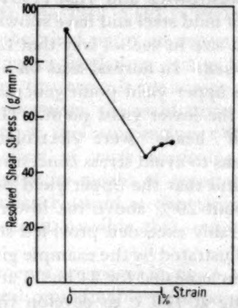


Fig. 3. Stress–strain curve showing an upper yield point 100% greater than the lower yield stress.

A further interesting feature shown by curve 1, fig. 2 is that plastic deformation occurred in the specimen before yielding took place, as is indicated by the deviation of the stress–strain curve from linearity below the stress-level corresponding to the upper yield point. Thus a yield point may appear in a specimen even though free dislocations have been present at some earlier stage during testing. Consider the deformation of an aged specimen in which some of the dislocations are pulled from their atmospheres at a value of the mean applied stress less than the upper yield stress under the action of stress-concentrators such as slight localized bending. These free dislocations will move and produce plastic deformation of the specimen but rapid work hardening will occur so that the free dislocations will become progressively more difficult to move. When work hardening has increased the apparent elastic limit of the 'free dislocation regions' to a value equivalent to the upper yield point, the remaining anchored dislocations will be pulled from their atmospheres and a yield point will be shown. This explanation requires that the number of free dislocations produced by stress-concentration effects should be small and should be localized in small regions of the crystal. Considerable numbers of widely dispersed, free dislocations would cause sufficient elastic disturbance of the lattice as they moved to pull the remaining anchored dislocations from their atmospheres at a stress well below the upper yield stress so that no significant yield point would appear.

§ 5. MAGNITUDE OF THE YIELD POINT

Sylwestrowicz and Hall (1951) have recently studied the deformation and ageing of mild steel and have shown that the magnitude of the yield point depends on grain size in such a way that the yield point values decrease as the grain size is increased. In normal tests on fine-grained mild steel these investigators found that the upper yield point generally appeared at a stress level some 60% above that of the lower yield point. However, this figure could be raised to almost 100% if 'heads' were electrolytically deposited on the polycrystalline wire specimens to avoid stress concentrations at the grips. In the case of zinc crystals it is found that the upper yield point generally occurs at a stress level which is only about 20% above the lower yield point value. However, this could be considerably exceeded provided stress-concentration effects were minimized and this is illustrated by the example given in fig. 3 which was taken after the specimen had been annealed for 24 hours at 50° C following several cycles of straining and annealing at 180° C to develop the yield point fully. The difference between upper and lower yield points in this case was 100% of the lower yield stress which proves that yield points can be produced in zinc crystals which are as marked as any which have been obtained in fine-grained mild steel.

§ 6. DEVELOPMENT OF THE YIELD POINT ON AGEING AT CONSTANT TEMPERATURE

An investigation into the development of the yield point in zinc at room temperature (Wain and Cottrell 1950) has shown that recovery occurs to a certain extent before the yield point appears. A similar phenomenon is illustrated in fig. 4 for a specimen aged at 50° C in which the recovery is revealed by a progressive decrease in the critical shear stress in curves 4 to 9. Once the yield point appears, further increases in the ageing time between successive tests causes a progressive increase in the yield point values (curves 10 to 12) due to more effective anchoring of the dislocations. This increase is linearly proportional to $\ln t$ where t is the ageing time (fig. 5) which is interesting since it is found (Churchman, private communication) that the change in electrical resistance produced as the carbon atoms move to dislocations during the strain ageing of iron also varies linearly with the logarithm of the ageing time.

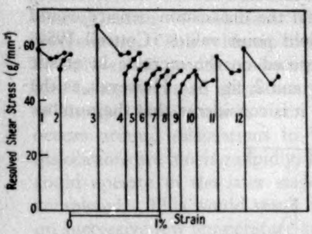


Fig. 4. Recovery followed by development of the yield point on ageing at 50° c.

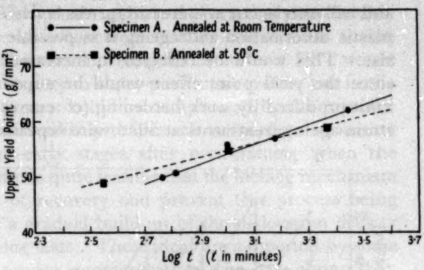


Fig. 5. Relationship between yield point value and ageing time at constant temperature.

Curve No.	1*	2	3	4	5	6	7	8	9	10	11	12
Time of ageing (hours)	1	0	20	0·5	1	1·5	2	2·5	3	6	17·25	39

* Specimen annealed at 180° c.

§7. EFFECT OF WORK HARDENING ON THE YIELD POINT

An important point in the behaviour of the yield point on repeated strain ageing is whether or not an accumulation of work hardening occurs from test to test. Recovery is believed to be complete in zinc crystals containing nitrogen on annealing at 180° c since, for a particular specimen, the yield stress remains constant on repeated strain-ageing treatments at this temperature (Wain and Cottrell 1950). An attempt to study the effect of work hardening on the development of the yield point in zinc was made by means of repeated strain-ageing tests at 50° c. Strain ageing is conveniently rapid at this temperature, but the temperature was low enough to make it possible that recovery was incomplete between tests.

Typical results of such a series of experiments are given in fig. 6. Strain-ageing treatments at 180° c were first given until reproducible yield point values were obtained (curve 1). The specimen was then annealed at 50° c for 24 hours (curve 2) and this annealing and testing cycle was repeated several times (curves 3 to 7). It is apparent that the yield point values decreased as testing proceeded and ultimately vanished. Annealing at 180° c caused the yield point to return at values characteristic of this annealing treatment (curves 8 to 10) and subsequent annealing for 24 hours at 50° c (curve 11) produced a yield point practically identical to that of curve 2.

It will be noted that, in both the tests taken after the specimen had been first annealed at 50° c following annealing at 180° c (curves 2 and 11), the yield points occurred at higher stress levels than those characteristic of annealing at 180° c. In the particular tests described, the specimen was tested at the annealing temperature; but similar phenomena were observed if the specimen was cooled to room temperature after each annealing at 50° c before the test was taken.

The most promising explanation of these effects on the basis of Cottrell's theory seems to be as follows. During yielding, and subsequently, the specimens are plastically deformed so that some work hardening occurs. In terms of dislocations this involves the multiplication of dislocations, thus increasing the dislocation density, and the passage of dislocations through the crystal (plastic deformation) becomes impeded by the stress fields of these extra dislocations and will only occur at increasing stress levels. Hence in a series of tests involving plastic deformation and ageing it is possible that the dislocation density would rise. This would be reflected in increased yield point values (Cottrell 1950) since the yield point effect would be superimposed on the increase in elastic limit produced by work hardening (cf. curves 1 and 2, fig. 6). However, as the strain-ageing treatments at 50° c were repeated, it is considered that the number of dislocations became so high due to plastic flow that they could no longer be completely anchored by the available nitrogen atoms. The yield point values would thus be expected to decrease until ultimately a smooth stress–strain curve would be obtained (curves 2 to 7, fig. 6). Subsequent annealing at 180° c produces complete recovery which means that the dislocation density reverts to the value obtaining before the strain-ageing tests at 50° c were commenced, and so a yield point is once again shown (curves 8 to 10, fig. 6). Support is provided for this explanation by the results of § 3 which show that, while the amount of nitrogen in these specimens is probably sufficient to satisfy the theoretical requirements as regards atmosphere formation around all dislocations when the specimen is in the annealed condition, it is insufficient when the dislocation density has risen to an extent which is below the maximum generally ascribed to heavily cold-worked metals.

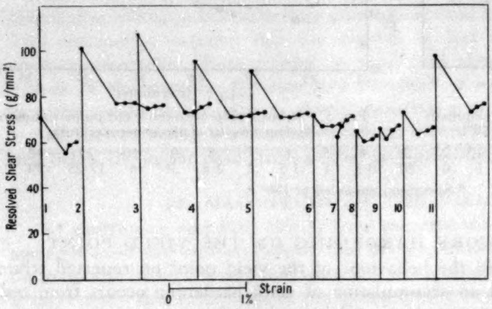


Fig. 6.

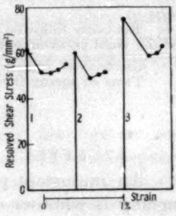


Fig. 7.

Fig. 6. Repeated strain-ageing experiments at 50° c.

Curve 1.	After annealing 1 hour at 180° c.
Curves 2 to 7.	After annealing 24 hours at 50° c.
Curves 8 to 10.	After annealing 1 hour at 180° c.
Curve 11.	After annealing 24 hours at 50° c.

Fig. 7. Stress–strain curves after various annealing treatments at 50° c and 180° c.

Curve 1.	After annealing for 1 hour at 180° c and cooling to room temperature.
Curve 2.	After annealing for 1 hour at 180° c, cooling to 50° c and maintaining at 50° c for 24 hours.
Curve 3.	After annealing for 24 hours at 50° c.

One obvious condition is necessary for the above explanation to be acceptable, namely, that recovery should not be complete in specimens aged for 24 hours at 50° c. Pure zinc is known to recover fairly rapidly even at room temperature, and the results presented in § 6 show that some recovery occurs on ageing at 50° c before a yield point starts to appear. Cottrell and Aytekin (1950) carried out recovery tests on high purity zinc single crystals and have shown that recovery, after a shear strain $\epsilon = 0{\cdot}05$ approximately, had progressed considerably after annealing 40 minutes at 60° c and was complete after 160 minutes at that temperature. This evidence indicates that annealing for 24 hours at 50° c would be sufficient to produce complete recovery in super-pure zinc crystals, but in the case being discussed the recovery process could be complicated by the presence of nitrogen in solution in zinc. It is unlikely that much atmosphere formation occurs around dislocations in the early stages after overstraining when the dislocations are moving rapidly, but it is quite feasible that the locking mechanism could operate in the later stages of recovery and prevent this process being completed. This would result in a gradual build-up of the dislocation density on successive low temperature ageing tests. These ideas are supported by some work of Cottrell and Dumbleton (private communication) who have shown that, in zinc crystals containing nitrogen in solution, strain ageing occurs when the rate of steady creep (which is equivalent to recovery) falls below a certain value, which depends on the temperature. These workers believe that, when creep or recovery is sufficiently slow, the dislocations wait so long at barriers, before making the next 'jump', that the nitrogen atoms have time to anchor them and they cannot make the 'jump' at all. This means that Cottrell and Aytekin's recovery results are not appropriate for zinc crystals containing nitrogen.

An alternative explanation of the fact that the yield point may disappear on repeated ageing tests at 50° c, which can be suggested on the basis of Cottrell's theory, is that the nitrogen content becomes depleted in the region of the slip planes during the progress of the tests. Since the activation energy for diffusion of nitrogen in zinc would be expected to be greater normal to the slip planes than parallel to them, strain ageing in zinc at low temperatures probably consists essentially of movement of nitrogen atoms along slip planes to dislocations. Hence, if the number of nitrogen atoms in the slip planes becomes gradually less during a series of strain-ageing experiments, the yield point values would be expected to decrease and ultimately vanish. Annealing at 180° c could provide the greater activation energy necessary for diffusion normal to the slip planes so that nitrogen atoms could be drawn from the body of the crystal to the slip planes and cause the yield point to return. There are two possible ways in which the nitrogen content could become depleted during repeated strain-ageing experiments. The first is that nitrogen atoms could become forced out of the specimen at the surface during deformation, and the second is that they could become permanently entrapped in vacant lattice sites which are believed to be produced by dislocation movement (Seitz 1950). In either case they would no longer be available to form atmospheres around dislocations and lock them in position. However, this explanation of the disappearance of the yield point is considered to be less satisfactory than that based on the accumulation of work hardening. In the first place there is no other evidence suggesting that depletion of the nitrogen atoms in the neighbourhood of slip planes occurs by either of the mechanisms outlined above. Also, on this basis, it is difficult to explain the observation that the yield point was developed at a higher stress level on annealing at 50° c than on annealing at 180° c, although this follows from the first explanation of the effect.

§ 8. EFFECT OF SURFACE CONDITION

A possible alternative explanation to that given in § 7 for the fact that the yield point obtained after annealing at 50° c was greater than that developed at 180° c is that it was related to the surface condition of the specimen. Harper and Cottrell (1950) have observed that oxide films on the surface of zinc specimens produce a considerable increase in the critical shear stress. It is possible that the paraffin oil in which the specimens were annealed could have oxidized the specimens at 50° c and reduced the oxide skin at 180° c since the oil has a tendency to 'crack' at high temperatures with the formation of free carbon. The results of experiments to test this point are given in fig. 7. The specimen was given several strain-ageing treatments so that the yield point was clearly developed, and was then annealed for 1 hour at 180° c, cooled to room temperature and tested giving curve 1. Curve 2 was obtained after the specimen had been annealed at 180° c, cooled to 50° c and maintained at that temperature for 24 hours, whilst curve 3 was the result of annealing for 24 hours at 50° c after overstraining. These results show quite clearly that the same yield point was developed after annealing at 180° c irrespective of any subsequent annealing at 50° c, so that the higher yield point on ageing at 50° c could not be a result of surface oxidation. The only difference between curves 2 and 3 was that for curve 3 the annealing treatment at 50° c followed overstraining whereas for curve 2 it followed annealing at 180° c, which proves that the higher yield point developed at 50° c is connected with the plastic deformation of the specimen. Confirmatory evidence that the above effect was not the result of surface oxidation was obtained from the fact that a specimen which had been electropolished also showed higher yield points on annealing at 50° c than were obtained after annealing at 180° c. Harper and Cottrell have shown that electropolishing renders the surface of zinc crystals very resistant to oxidation so that no change in critical shear stress was observed if an electropolished specimen was exposed to steam for 1 hour, although this treatment produced a marked increase in the critical shear stress of unpolished specimens.

§ 9. DISCUSSION

Nitrogen absorption, of the order of 0·002 wt. % (0·014 atomic %), has been shown to occur in zinc treated in such a way as to produce a yield point, and microscopic examination has indicated that this nitrogen is present in solid solution. These results, taken in conjunction with previous work, leave little doubt that the yield point in zinc is due to dissolved nitrogen.

The theoretical prediction that a yield point should appear in a susceptible specimen without any prior deformation is verified by experiment in the case of zinc. However, zinc crystals are generally extremely weak so that accidental bending during handling frequently gives rise to stress-concentration effects in the first few tests which cause the development of the yield point to be delayed. Arguments have been advanced to explain the observation that a yield point may be shown even though a small number of free dislocations has been present in the specimen at some stage of the test. Similar arguments can account for the fact that light filing of the surface of an aged iron specimen, which would produce small work hardened regions, has little influence on the yield point values (Sylwestrowicz and Hall 1951).

Yield points have been developed in zinc crystals which are relatively as large as any obtained in fine-grained mild steel. This is further evidence that the yield point phenomenon in different metals is essentially the same and can be expected to be produced by similar mechanisms.

The results presented enable a clearer picture to be obtained of the effect of strain hardening on the development of the yield point. In a series of tests involving plastic deformation and ageing the dislocation density will rise provided the ageing temperature is not so high as to produce complete recovery before ageing commences. Theoretically this can have one of two consequences. In the first place, provided sufficient solute atoms are available to lock all the dislocations produced, the yield point will return at higher levels of stress after each successive ageing treatment since the yield point effect will be superimposed on the normal increase in elastic limit due to work hardening (Cottrell 1950). This has been observed during the strain ageing of polycrystalline iron (Muir 1906) and iron single crystals (Cottrell 1950, Churchman and Cottrell 1951). However, it is possible that, where the solute concentration is low, there may be sufficient solute atoms to anchor all the dislocations in an annealed crystal, but insufficient to form effective atmospheres around the increased number produced during plastic flow. This appears to apply to the zinc crystals used in this investigation. Under such conditions there will be two opposing influences on the yield point during repeated strain-ageing tests: a tendency to increase the yield point due to accumulated work hardening, and a tendency to decrease the yield point due to the dislocations becoming less effectively anchored as the number of solute atoms available to each individual dislocation becomes progressively less. This latter process is, in effect, the reverse of that which produced the increase of the yield point with increasing ageing time at constant temperature as discussed in § 6. According to this theory, zinc would be expected to behave in a similar manner to iron, rather than in the way described above, provided the amount of nitrogen in solid solution could be sufficiently increased. This hypothesis is being tested by further experiment.

ACKNOWLEDGMENTS

This work forms part of the research programme of the Aeronautical Research Laboratories, Department of Supply, Melbourne, Australia, and thanks are due to the Chief Scientist, Department of Supply, for permission to publish. The author is indebted to Professor A. H. Cottrell, Metallurgy Department, University of Birmingham for his continued interest in the work, to Mr. S. Martin of the Defence Research Laboratories, Department of Supply, who made the gas analyses, and to Mr. C. B. Rogers for technical assistance.

REFERENCES

ANDRADE, E. N. DA C., and ROSCOE, R., 1937, *Proc. Phys. Soc.*, **49**, 152.

BARRETT, C. S., and LEVENSON, L. H., 1940, *Trans. Amer. Inst. Min. Met. Engrs.*, **137**, 76, 211.

CHURCHMAN, A. T., and COTTRELL, A. H., 1951, *Nature, Lond.*, **167**, 943.

COTTRELL, A. H., 1948, *Report on the Strength of Solids* (London : Physical Society) p. 30; 1949, *Progress in Metal Physics* I (London: Butterworths Scientific Publications) p. 96; 1950, *Report of Pittsburgh Conference on Plastic Deformation of Crystals* (Washington : Carnegie Institute of Technology and Office of Naval Research) p. 60.

Cottrell, A. H., and Aytekin, V., 1950, *J. Inst. Met.*, **77**, 389.
Cottrell, A. H., and Bilby, B. A., 1949, *Proc. Phys. Soc.* A, **62**, 49.
Cottrell, A. H., and Gibbons, D. F., 1948, *Nature, Lond.*, **162**, 488.
Harper, S., and Cottrell, A. H., 1950, *Proc. Phys. Soc.* B, **63**, 331.
Muir, J., 1906, *Proc. Roy. Soc.* A, **77**, 277.
Nabarro, F. R. N., 1948, *Report on the Strength of Solids* (London : Physical Society) p. 38.
Orowan, E., 1934, *Z. Phys.*, **89**, 634; 1940, *Proc. Phys. Soc.*, **52**, 14.
Polanyi, M., 1925, *Z. Tech. Phys.*, **6**, 121.
Schwartzbart, H., and Low, J. R., 1949, *J. Metals*, **1**, 637, 877.
Seitz, F., 1950, *Phys. Rev.*, **79**, 1002.
Smith, C. L., 1947, *Nature, Lond.*, **160**, 466.
Sylwestrowicz, W., and Hall, E. O., 1951, *Proc. Phys. Soc.* B, **64**, 495.
Wain, H. L., and Cottrell, A. H., 1950, *Proc. Phys. Soc.* B, **63**, 339.

A Critical Study of the Asymmetrical Temperature Gradient Thermoelectric Effect in Copper and Platinum

By N. FUSCHILLO
University of Leeds

Communicated by E. C. Stoner; MS. received 8th July 1952

ABSTRACT. A method is described by which metallic wires can be subjected to easily measurable but very pronounced asymmetrical temperature gradients without straining the wire. The electromotive forces developed by the application of symmetrical and asymmetrical gradients to annealed 'thermopure' platinum wire and specially prepared cables of copper wire, at both high and low temperatures, have been measured at many points along considerable lengths of wire. Precautions to minimize or eliminate spurious effects enabled measurements significant to 0·01 μv to be taken.

The results provide no justification for the supposition that an asymmetrical temperature gradient thermoelectric effect exists, under the conditions of investigation, which cannot be explained in terms of inhomogeneity effects. The nature of these effects is discussed and some evidence is provided which indicates that the variation in the Thomson thermoelectric potential with strain and impurity is significant.

§ 1. INTRODUCTION

It has been known for many years that the application of an asymmetrical temperature gradient to an apparently homogeneous metallic wire results in the production of an electromotive force. There is, however, little agreement between the results obtained by different observers, and some of the hypotheses which have been used to interpret the results are far from satisfactory. It is clearly desirable that systematic investigations should be carried out in an attempt to determine whether the effects which have been reported arise from fundamental characteristics of metals, or are of secondary 'experimental' origin. Apart from this the results are of considerable practical interest in connection with precision thermoelectric thermometry. Indeed, the necessity for the detailed investigation recorded here became apparent during the development of a new design for a standard low temperature thermo-element and an improved method of testing and selecting wire to be used in the construction of thermoelectric thermometers.

The purpose of the experiments was to provide a rigorous test of the existence of the effect by correlating the effects produced by symmetrical gradients with those produced by gradients of pronounced asymmetry at both high and low temperatures and at many points along considerable lengths of wire. Attention has been paid to the elimination of strain, impurity and hidden sources of error. Apart from this a method is introduced of producing gradients of greater asymmetry than those hitherto used, and a means of reducing the inhomogeneity effects in copper wires is described.

§ 2. PREVIOUS WORK

The experimental method invariably used in previous work has been to apply an asymmetrical temperature gradient along a metallic specimen, the ends of which were maintained at the same temperature: the electromotive force developed was then measured with either a galvanometer or a potentiometer. Large but unmeasurable asymmetrical gradients have been produced by heating a bar of metal on one side of a constriction machined in it, but this method is unsuitable because of the large stresses that can be set up in the constriction due to thermal expansion (Benade 1921). On the other hand small but easily measurable asymmetrical gradients have been produced by suspending a wire along the horizontal axis of a non-uniformly wound electric heater (Feng and Band 1934, Ch'en and Band 1939, Fraser 1938). In this method, as used by Band and his co-workers, the asymmetrical temperature gradient ratio was 1 : 6, i.e. the experimental arrangements were such that six times more wire was included in the less steep gradient than in the steep gradient, the distances being measured from the point of maximum temperature to the points at which the temperature of the specimen was not significantly different from that of the surroundings. Measurements were made apparently at only one relative position of the heater and wire, which was neither annealed nor particularly pure. Fraser (1938), using the wire method with commercial copper, but with an asymmetrical gradient ratio of only 1 : 2 and a maximum temperature of 300° c, was the first to show that large variations in the electromotive force were obtained at different points along the wire. He observed that annealing the wire reduced the order of the electromotive force produced from 1·5 to 0·4 μv, whilst an annealed copper wire drawn from spectroscopically pure rod gave a maximum e.m.f. of about 0·07 μv and an average value, irrespective of sign, of almost 0·04 μv. Hence Fraser came to the conclusion that the large effects observed by Band and others were due to a slow unequal annealing of the wire in the temperature gradients and that any residual effects after annealing were due to a non-uniform distribution of impurity in the wire. Ch'en and Band (1939) considered the annealing hypothesis difficult to maintain and proposed and examined in some detail a positive emission hypothesis with which they attempted to explain not only their own results but those of other workers, including Fraser.

It is remarkable that in most of the previous work no investigations into the effects of symmetrical temperature gradients have been made. Such investigations lead naturally to an estimate of the importance of factors such as local inhomogeneities. In addition, the limitation of the temperature range to 0° c to 300° c does not enable generalizations to be made from the results with confidence. The restriction of the measurements to one position in the specimen, except for the work of Fraser (1938), and the small degrees of asymmetry employed, constitute severe restrictions of the experimental methods.

The proposed explanations of the observed effects may be divided into two distinct groups as follows:

(*a*) That the e.m.f's produced by asymmetrical temperature gradients constitute a new thermoelectric effect, not explicable on the classical theory of Kelvin (Benedicks 1920). The success of some workers in correlating the observed e.m.f. with the temperature distribution along the wire by means of a quadratic formula (Feng and Band 1934, Pi and Band 1935, Li and Band 1935, Band 1935, Ch'en and Band 1935) led to the suggestion that there exists a component of the Thomson thermoelectric potential dependent on the asymmetry of the temperature gradient.

(*b*) That the effect is essentially spurious, being due to one of the following: (i) hidden sources of error or unsuspected heterogeneities due to strain or impurities (Benade 1921, Young 1925, Terada, Tsutsui and Tamano 1928, Fraser 1938); (ii) a slow non-uniform annealing taking place in the asymmetrical temperature gradients: the annealed wire is thermoelectrically different from the less annealed wire, and a Seebeck e.m.f. is generated (Fraser 1938); (iii) a disturbance of the Thomson thermoelectric potential caused by unequal positive emission from the surface of the specimen due to the asymmetrical temperature distribution (Ch'en and Band 1939); or (iv) thermoelectric anisotropy in metals which do not crystallize in a cubic form (Bridgman 1934).

§ 3. PRESENT WORK

The method employed to produce asymmetrical gradients differed from those used by other workers in that the less steep gradient was produced by winding the wire without strain on an insulated metallic former dipping vertically into a bath of liquid nitrogen or hot oil contained in a Dewar vessel. The steep gradient was produced by taking the other end of the wire vertically out of the liquid, maintaining it in thermal contact with, but electrically insulated from, the coils. The arrangement as used for platinum is shown in fig. 1. The degree of asymmetry which can be produced by this method is limited only by the amount of wire available and the electrical resistance of the circuit, for copper asymmetries of 1000 : 1 could be produced if desired. The temperature distribution as a function of length along the wire can be measured easily and accurately by suitably placed thermocouples along the metallic former. The relative positions of the wire and gradient are easily altered, independently of each other or otherwise, by altering the depth of immersion of the coil and adjusting the position of the straight wire LH. This arrangement has the additional advantage that various symmetrical gradient tests can be applied by immersing the slack loop HIJ only. As there is good thermal contact between the wire, the former and the bath liquid, the rapidity with which measurements may be made is considerably increased. Also the wire can be annealed *in situ* on the former if desired. A disadvantage of the method is that the symmetrical gradient tests over the wire used in the coils can be performed only before coiling.

Experiments were performed on annealed 'thermo-pure' platinum wire and specially prepared cables of high grade insulated copper wire. These materials were used because of their immediate availability and their relevance to thermoelectric thermometry. The crystalline form is face-centred-cubic in each case, so that these materials should show no thermoelectric anisotropy.

To assess the degree of inhomogeneity and the ease with which it could be increased by cold working and tension, in the copper and platinum wire used, loop tests (Giauque, Buffington and Shulze 1927) were performed before and after the wires were subjected to the above treatments. These tests involved the passing of a loop of wire through liquid nitrogen and measuring the e.m.f. produced by the symmetrical gradients at many points along the wire. The measurements showed that the mechanical treatment the wire undergoes in the final experiments was not sufficient to affect the results significantly.

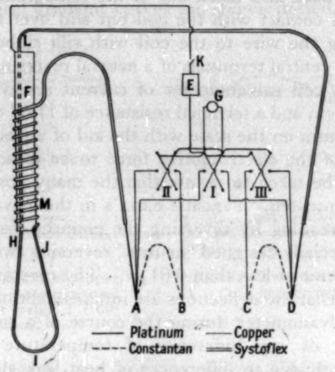


Fig. 1. Experimental arrangement for platinum, giving asymmetrical temperature gradient ratio 1 : 10. The insulated wire HL is in sliding contact with the inner surface of the former and can be pulled out at the point L.

(i) *Experiments with Copper Wire*

Because of the high values of the electromotive forces produced on loop testing single wires the effects of inhomogeneities were reduced in the high grade insulated copper wires used by grouping many wires in parallel and wrapping them with silk thread. This procedure resulted in a more uniform distribution of inhomogeneity, as may be seen from curves 1 and 2 of fig. 2; these curves

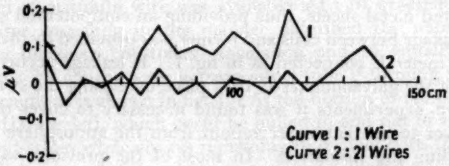


Fig. 2.

represent the resultant electromotive force obtained on increasing the depth of immersion into liquid nitrogen of a single 36 SWG copper wire loop and of a loop consisting of 21 such wires in parallel. For the final experiments one set of measurements was performed on each of two 12-metre cables, one consisting of forty 40 SWG and the other of twenty-six 36 SWG chemically covered wires.

The first was the more successful: inhomogeneity effects due to the low temperature symmetrical gradients obtained on loop testing were reduced to 0·01 μv.

The 40 SWG cable was wound without tension on a brass tube of length 30 cm, diameter 4·3 cm and wall thickness 0·3 mm to form a single-layered coil of length 14 cm. A preliminary experiment had shown that the differential contraction of the wire and the former on cooling to liquid nitrogen temperature did not introduce any significant strain inhomogeneity. The wire from the bottom of the coil was looped back and taken out of the temperature gradient so that it was in good thermal contact with the coil but slid over it very easily. It was secured by wrapping the wire to the coil with silk ribbon. The ends of the cable were led to the central terminals of a neutral reversing switch to which was connected a moving coil galvanometer of current sensivity 21·70 cm/μA at a scale distance of 240 cm and a terminal resistance of 11·67 ohm. The deflections could be read to 0·1 mm on the scale with the aid of a reading lens.

The magnitude of the electromotive force to be measured is so small that particular care must be taken to ensure that the many possible sources of error are minimized or eliminated. Parasitic e.m.f's in the galvanometer circuit were eliminated at every reading by reversing the connections of the galvanometer to the cable with a specially designed 'neutral' reversing switch which introduced an e.m.f. into the circuit of less than 0·01 μv. The reversing procedure has the additional advantage that the deflections are independent of any zero drift which takes place in the galvanometer during the course of a run. To eliminate the thermoelectric effect of any difference of temperature between the central terminals of the switch due to differences of heat leak along the two portions of the cable a differentially connected copper–constantan thermocouple with insulated junctions was screwed into the central terminals with the copper cable. The leads from this thermocouple could be led to the galvanometer through another 'neutral' reversing switch. Direct calibration in terms of the e.m.f's produced in the switches was effected by warming one of the central terminals slightly with a small heating coil and measuring the deflections on reversals of the switches due to the thermocouple and the switch successively when the rest of the apparatus was at room temperature. Tests made at regular intervals throughout the experiments showed that no heat-leak effect greater than 0·01 μv occurred. All parts of the apparatus were electrically shielded by placing them on interconnected metal sheets, thus providing an equipotential surface. A test for electrical leakage between coils and former was provided by the 24-volt probes of a resistance meter E connected as in fig. 1. If leakage occurred a change in the deflection of the galvanometer G took place on closing the switch K. In the low-temperature experiments it was found necessary to cover the coils with a thin plastic cover to prevent water vapour from the atmosphere condensing on them and spoiling the insulation. In most of the previous work insufficient consideration appears to have been given to possible sources of error.

Experiments were performed at temperatures of 190°C and −190°C by immersing various parts of the wire in a Dewar vessel containing well-stirred electrically heated oil or liquid nitrogen. There was no electrical interference from the oil heater circuit. The experimental procedure was as follows: Before and after the temperature gradients were applied the residual electromotive force in the circuit due to the reversing switch was measured by reversals and was found not to be greater than 0·005 μv. A slack loop one metre long was then formed at the bottom of the coils. This was pushed vertically into the liquid, the coil being at room temperature, and readings were taken at 5 cm intervals along the wire. This symmetrical gradient test gave a good indication of the homogeneity of the wire. Then 2 cm of the coil was immersed and the slack wire pulled out through the temperature gradient without straining the wire significantly. Readings were again taken at 5 cm intervals of length. In this way an asymmetric gradient ratio of 1 : 60 was applied. The experiments were repeated with the 36 SWG copper cable. The results are shown in fig. 3 and fig. 4. The symmetrical gradient curves are shown dotted.

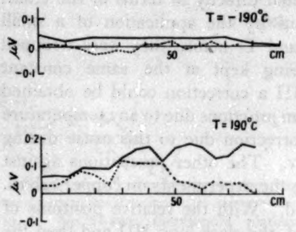


Fig. 3. Forty No. 40 SWG insulated copper wires in parallel. —— asymmetrical temperature gradient ratio 1 : 60; - - - - symmetrical gradient.

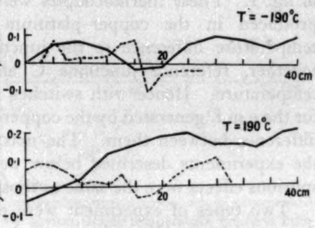


Fig. 4. Twenty-six No. 36 SWG insulated copper wires in parallel.

B
5
2

(ii) *Experiments with Platinum Wire*

Similar experiments were performed with 34 SWG 'bright drawn' 'thermopure' platinum wire. This was annealed for four hours by suspending a loop two metres long in a vertically mounted Pyrex tube evacuated to a pressure of 10^{-3} mm Hg, and passing an alternating current through the wire to raise the temperature to 1200° C. To give some indication of the extent to which annealing and cold working due to coiling affected the results, loop tests in liquid nitrogen (Giauque, Buffington and Shulze 1927) were performed. The average inhomogeneity e.m.f. irrespective of sign was reduced from 0·50 μv to 0·09 μv by the annealing process.

As only 4 m of platinum wire was available a 1 : 10 gradient of asymmetry was used. After annealing in the manner described the wire was threaded through 1 mm diameter 'Systoflex' insulation and 14 equally spaced turns were wound on a 2·0 cm diameter german silver tube F forming a single-layered coil 12 cm in length as is shown in fig. 1. At the bottom of the coil about 30 cm of slack loop HIJ was left, which could be pulled in and out of the temperature gradient without strain while always remaining in contact with the inner surface of the tube. To each end of the platinum wire was soldered a 36 SWG dsc copper wire which led to the central terminals of the neutral reversing switch I. As a temperature difference of 0·002° C between the platinum–copper junctions generates an e.m.f. of about 0·01 μv, care was taken that they should be maintained at the same temperature as closely as possible. To correct for any difference of temperature that might arise due to an unequal amount of heat leak from the platinum wires a copper–constantan thermocouple was soldered to each of the copper–platinum junctions. Hence altogether the four junctions ABCD are obtained as is shown in fig. 1. Junctions A and D were placed on either side of a strip of thin plastic adhesive tape and the wires leading to them covered over 30 cm with a layer of the same material. The plastic cover around the junctions was wrapped with three layers of copper wire forming a conducting sheath for the equalization of temperature around them. The copper–constantan junctions C and B were similarly treated. Finally the two pairs of junctions were wrapped together and immersed 25 cm below the surface of oil contained in a Dewar vessel. The connections of the copper–constantan thermocouples to the galvanometer through the neutral reversing switches II and III were as shown in fig. 1. These thermocouples were calibrated directly in terms of the e.m.f. produced in the copper–platinum junctions by the application of a small temperature difference to the junctions A and D before they were wrapped together, reference junctions C and B being kept at the same constant temperature. Hence with switches II and III a correction could be obtained for the e.m.f. generated by the copper–platinum junctions due to any temperature difference between them. The maximum correction due to this cause during the experiments described below was $0{\cdot}02\,\mu$v. The other precautions against spurious effects were the same as those used in the experiments on copper cables.

Two types of experiment were performed. With the relative positions of the straight wire HL and the coiled wire fixed, the slack loop HIJ and then the straight wire HL and the coil were lowered vertically into liquid nitrogen or hot oil at 190° c contained in a Dewar vessel. Readings were taken, as in the experiments on copper, but at 1 cm intervals as the depth of immersion was increased. In this way the electromotive forces generated by the symmetrical gradients over the slack loop and then by the asymmetrical gradients over the coiled wire and straight wire HL were measured at many points. The results are shown in fig. 5. The second experiment consisted of comparing a point J on the slack wire about 4 cm from the end of the coil with many points on the rest of the slack wire under symmetrical gradient conditions and then comparing the same points on the slack wire with a point M on the coil. The results are shown in fig. 6.

§4. DISCUSSION

For copper the average value of the e.m.f. for asymmetrical gradients is about two or three times the average value for symmetrical gradients at both high and low temperatures. It is significant that the larger the average value of the symmetrical gradient e.m.f. the larger is the average value of the asymmetrical gradient effect. Since the former can only be due to some form of inhomogeneity it appears inescapable that the latter must be related to it. As 60 times more wire is included in one side of the gradient than in the other we may expect a more thorough averaging of the inhomogeneity effects in the less steep gradient. Hence we may expect the curves referring to asymmetrical gradients to be smoother than those for the symmetrical gradients. Since the electromotive force developed by a given inhomogeneity will depend upon the temperature difference across it, we may also expect the average asymmetrical gradient e.m.f. to be different from that produced by the symmetrical gradients, since the temperature gradient over any of the more inhomogeneous sections of the wire will be less on the less steep side and hence the electromotive force produced will also be less. These expectations are realized in both sets of experiments, as may be seen from figs. 3 and 4. Such considerations are more applicable to the investigations of Band *et al.*, where a gradient of asymmetry 1 : 6 on unannealed and not specially pure wire was used (since symmetrical gradient effects should be large) than to the work of Fraser (1938) on annealed copper wire with a gradient of asymmetry 1 : 2. It may be mentioned that the asymmetrical gradient curves given by Fraser (1938) are similar to those obtained in this investigation for symmetrical gradients. Similar considerations apply to the experiments on platinum if allowance is made for the smaller asymmetrical ratio of the temperature gradients employed and the larger symmetrical gradient e.m.f's.

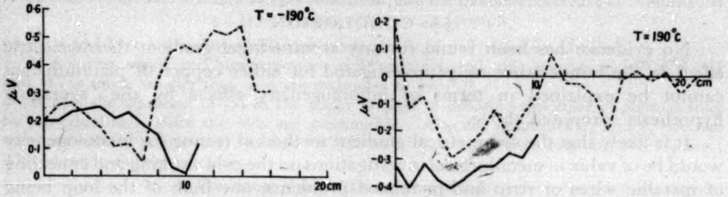


Fig. 5. No. 34 SWG 'thermopure' platinum wire. —— asymmetrical temperature gradient ratio 1 : 10; - - - - symmetrical gradient.

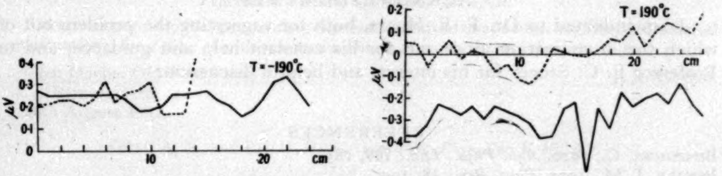


Fig. 6. No. 34 SWG 'thermopure' platinum wire. —— asymmetrical temperature gradient ratio 1 : 10; - - - - symmetrical gradient.

If the inhomogeneity hypothesis is correct there remains the question of the mechanism by which they produce the phenomena observed. The inhomogeneities are of two kinds, chemical and physical. The latter are either due to strain or to differences in crystalline structure at various places along the wire. The difference in the electromotive force produced by the same temperature gradients on different portions of the wire can be due either to an integrated effect of the Seebeck e.m.f's caused by the inhomogeneities or to the variation of the Thomson effect with strain and impurity. Certain features of the work described indicate that the mechanism last mentioned is likely to be significant. The directly measured value of the Thomson coefficient for copper is only $-0{\cdot}07\,\mu v/°c$ at $-130°c$ and $-0{\cdot}80\,\mu v/°c$ at $-70°c$, and increases linearly with increase in temperature, reaching a value of $-2{\cdot}4\,\mu v/°c$ at $+130°c$ (International Critical Tables 1929). With the temperature differences used, the Thomson e.m.f. generated is estimated to be about $160\,\mu v$ for the low temperature experiments and about $400\,\mu v$ for those at high temperatures. Hence we would expect a greater asymmetrical effect at high temperatures than at low. The measurements do in fact show that the ratios of the average values of the difference between the asymmetrical and symmetrical electromotive forces at high and low temperatures are 3·5 and 3·0 for the 40 SWG and 36 SWG copper cables respectively. Similar considerations apply to platinum, where the Thomson coefficient is much larger than for copper, but its variation with temperature is only a quarter of that exhibited by copper. Typical values are $9{\cdot}7\,\mu\mathrm{v}/^\circ\mathrm{c}$ at $-70^\circ\mathrm{c}$ and $9{\cdot}3\,\mu\mathrm{v}/^\circ\mathrm{c}$ at $+130^\circ\mathrm{c}$ (International Critical Tables 1929). Hence we may expect the shapes of the symmetrical and asymmetrical gradient curves to show a similar point-to-point variation in magnitude at high and low temperatures. Reference to fig. 5 and fig. 6 shows that this is approximately the case, especially in the symmetrical gradient curves.

§ 5. CONCLUSION

No evidence has been found for any asymmetrical gradient thermoelectric effect in the temperature range investigated for either copper or platinum that cannot be explained in terms of inhomogeneity effects by the 'averaging hypothesis' proposed above.

It is likely that the symmetrical gradient method of testing for inhomogeneity would be of value in metallurgical investigations on the cold working and annealing of metallic wires or strip and in related problems, one limb of the loop being unchanged for reference purposes.

ACKNOWLEDGMENTS

I am indebted to Dr. F. E. Hoare, both for suggesting the problem out of which this investigation arose and for his constant help and guidance, and to Professor E. C. Stoner for his interest and helpful discussion.

REFERENCES

BENEDICKS, C., 1920, *Ann. Phys., Lpz.*, **169**, 185.

BENADE, J. M., 1921, *Phys. Rev.*, **18**, 199.

BRIDGMAN, P. W., 1934, *The Thermodynamics of Electrical Phenomena in Metals* (London : Macmillan).

CH'EN, L. C., and BAND, W., 1935, *Proc. Phys. Soc.*, **47**, 904; 1939, *Proc. Camb. Phil. Soc.*, **35**, 622.

FENG, P. C., and BAND, W., 1934, *Proc. Phys. Soc.*, **46**, 515.

FRASER, M., 1938, *Phil. Mag.*, **25**, 1938.

GIAUQUE, W. F., BUFFINGTON, R. M., and SHULZE, W. A., 1927, *J. Amer. Chem. Soc.*, **49**, 2343.

International Critical Tables, 1929, **6**, 228 (London: McGraw-Hill).

LI, M. K., and BAND, W., 1935, *Proc. Phys. Soc.*, **47**, 859.

PI, T. H., and BAND, W., 1935, *Proc. Phys. Soc.*, **47**, 852.

TERADA, T., TSUTSUI, T., and TAMANO, M., 1927–28, *Sci. Pap. Inst. Phys. Chem. Res., Tokyo*, **7**, 20.

YOUNG, J., 1925, *Proc. Phys. Soc.*, **37**, 145.

LETTERS TO THE EDITOR

Experiments on the Mechanics of Rubber
II: The Torsion, Inflation and Extension of a Tube

On reading the paper by Gent and Rivlin (1952) of above title, it occurred to me that an interesting comparison with the volume change in a twisted rubber tube would be that of a 'tube' formed of a large number of thin rigid inextensible rods of equal length arranged parallel to each other in cylindrical formation, their ends being pivoted in ball-and-socket manner in two circular end-plates.

If one plate is rotated relative to the other through an angle ϕ, a section of the tube in the plane of the axis assumes a hyperbolic form, and the fractional decrease of 'volume' is

$$1-\tfrac{1}{3}(2+\cos\phi)\{1-2a^2(1-\cos\phi)/l^2\}^{1/2},$$

where a and l are the radius and length respectively of the undeformed tube.

If ϕ is very small, and higher powers of ϕ^2 are neglected, this reduces to $\frac{1}{2}\phi^2(a^2/l^2+\frac{1}{3})$ and, if l is much smaller than a, it approximates to $\frac{1}{2}\phi^2a^2/l^2$.

If now the equation (2.15) in the paper is used to describe this tube of rods, λ may be put equal to 1, since the rods are inextensible. Also the quantities I_1 and I_2 become zero and infinity respectively, so that $\partial W/\partial I_1$ may be considered infinite and $\partial W/\partial I_2$ zero. The equation thus reduces to

$$\text{Fractional Volume Decrease}=\tfrac{1}{2}\psi^2a_1^2.$$

But equation (2.15) was obtained on the assumption that end-effects are negligible. If therefore $\psi=\phi/l$, where ϕ is the angle of torsion over a short length l at the centre of the tube, we have

$$\text{Fractional Volume Decrease}=\tfrac{1}{2}\phi^2a_1^2/l^2.$$

This corresponds to the approximate expression derived above from ordinary geometrical considerations, for a tube of rods where the length is much shorter than the radius.

Glen Doone, Old Fold View,
Barnet, Herts.
26th August 1952.

J. R. Gould.

Gent, A. N., and Rivlin, R. S., 1952, *Proc. Phys. Soc.* B, **65**, 487.

The Growth and Orientation of Single Crystals of Indium

Single crystal rods of indium, about 3 mm in diameter and up to 15 cm in length, have been grown from the melt using the method of Andrade and Roscoe (1937). The satisfactory growth depends critically on the purity of the metal. At first 99·9+% pure indium was used, but with varying degrees of success. More reliable results were obtained after the indium had been electrolytically purified using the method of Baxter and Alter (1933). The electrolytic indium was deposited in 'tree' form on the cathode, from which it could be easily removed. The purity of the metal was not determined accurately, but it is believed to contain less lead, iron and nickel than the 99·9+% indium. The amounts of the insoluble impurities iron and nickel seem to be of particular importance for the successful growth of the single crystals.

The metal, before crystallization, was cast *in vacuo* in a Pyrex tube coated internally with a graphite suspension. The cast rod was then removed from its envelope and was placed in a 4 mm bore Pyrex tube coated internally with the graphite suspension to prevent attack of the glass by the indium. The Pyrex tube was inserted in the horizontal furnace tube of the crystallization apparatus and the air pressure inside the furnace tube was reduced to approximately 10 mm Hg. The travelling furnace was then run over the specimen rod at the rate of 0·5 mm/min and the furnace current adjusted so that the maximum temperature reached by the metal was about 175°c. Under these conditions a thin oxide film was formed over the surface of the rod; this prevented the molten rod breaking up into sessile drops.

.B
5
2

After crystallization the rod was electropolished using a solution of 1 part concentrated nitric acid to 2 parts methyl alcohol (Carapella and Peretti 1949). A current density of $30\ \mathrm{A/dm^2}$ was found to be satisfactory. The electropolishing removed the oxide film and any remaining graphite and produced a bright metal surface on which sharp etch pits could be produced by the subsequent chemical etching.

The various chemical etchants suggested or used by previous investigators for indium were found to be unsatisfactory. With Villela's reagent, recommended by Carapella and Peretti (1949), though the crystal boundaries were revealed, the optical reflection pattern from the etch pits was too blurred for an accurate determination of the crystallographic orientation. Etchants consisting of hydrochloric acid alone, or alternate immersion in hydrochloric acid and dilute nitric acid (Gwathmey, private communication), or 50% nitric acid alone (Carpenter and Tamura 1927), were equally unsuccessful. A new etchant, which has proved eminently suitable, was therefore developed. The crystal was immersed in concentrated hydrochloric acid to which sufficient potassium chlorate had been added to provide an abundant supply of chlorine during the etching process, which lasted about 5 minutes. (A similar etchant was used by McKeehan and Hoge (1935), though in their case for tin and not indium.)

The crystallographic orientation was determined by the optical etch-pit reflection method using the optical goniometer designed by Goss (1952). The reflection patterns consisted principally of reflections from the (001), (010) and (111) planes. For the determination of the orientation, i.e. the angle ψ between the rod axis and the c-axis of the crystal, using a plot of the reflection pattern angles on a Wulff net, the angles between some of the crystallographic planes of indium were calculated (see table). Unit cell dimensions of $a=4{\cdot}588$ Å and $c=4{\cdot}938$ Å (Wyckoff 1951) were used. It should be noted that since $c/a=1{\cdot}076$ the structure is almost cubic, and thus accurate measurements of the reflection pattern are essential in order that the (001) pole may be unmistakably picked out. In practice, angles in agreement to within 1 degree of those given in the table were obtained. For the various single crystals grown the values of ψ were all betweeen 50 and 90 degrees.

Table of Crystallographic Angles of the Indium Structure

Planes	(101)	(100)	(111)	(110)
(001)	47°	90°	57°	90°
(111)	36°	54°	—	33°
(011)	62°	90°	36°	59°

The single crystals were extremely ductile and had to be handled with great care in order to avoid deformation and twinning. The crystals are being used to determine the variation with temperature of the linear thermal expansion coefficients of indium.

The authors wish to thank Mr. S. Weintroub for his encouragement and advice and Professor A. M. Taylor for laboratory facilities. One of them (A. J. G.) is indebted to the Ministry of Education for a Further Education and Training Grant and the other (E. V. V.) to the University authorities for a research grant.

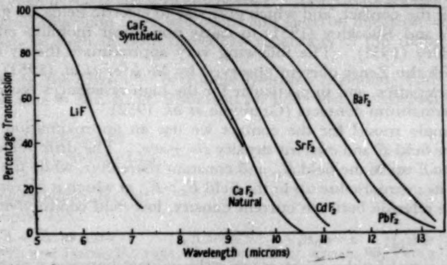
The Physical Laboratory,
University of Southampton.
15th August 1952.

A. J. Goss.
E. V. Vernon.

Andrade, E. N. da C., and Roscoe, R., 1937, *Proc. Phys. Soc.*, **49**, 152.
Baxter, G. P., and Alter, C. M., 1933, *J. Amer. Chem. Soc.*, **55**, 1943.
Carapella, S. C., and Peretti, E. A., 1949, *Metal Progress*, **56**, 666.
Carpenter, H. C. H., and Tamura, S., 1927, *Proc. Roy. Soc.* A, **113**, 161.
Goss, A. J., 1952, *Thesis*, University of London.
McKeehan, L. W., and Hoge, H. J., 1935, *Z. Kristallogr.*, **92**, 476.
Wyckoff, R. W. G., 1951, *Crystal Structures*, Vol. 1 (New York: Interscience Publishers), Chap. II f.

Infra-Red Transmission Limits of Single Crystals of some Fluorides

We have grown single crystals of the fluorides of strontium, barium, cadmium and lead, in sizes up to 6 cm in diameter and up to 800 g in weight. It should be possible to make much larger crystals without undue difficulty, and our present specimens are intended only for an investigation of the optical properties of these materials. The figure shows their transmissions in the infra-red region, after reflection losses have been eliminated; the observations have been reduced to a standard thickness of 1 centimetre for each material, and the absorption coefficients can therefore be determined directly from the figure.



The transmissions of lithium and calcium fluorides have been plotted on the same diagram to show how the newer materials compare with them: it may be worthy of note that the curve labelled 'CaF_2 synthetic' applies to a single crystal that we have grown from chemically produced calcium fluoride powder, and not from the scrap natural material of optical quality which had previously been considered a prerequisite for successful growing of artificial fluorite. The slight difference between the transmissions of natural and synthetic fluorites as shown in the diagram may well be due to the errors involved in reducing the observations on the synthetic specimen, which was 50 mm thick, and on the natural specimen, which was 4 mm thick.

The four new fluorides all transmit further into the infra-red than fluorite, while they all have good water-resisting properties. The approximate refractive indices for yellow light, and the 50% transmission limits in the infra-red, are as follows:

	Ref. ind.	50% limit
Strontium fluoride	1·43	10·7 microns
Barium fluoride	1·47	11·7 ,,
Cadmium fluoride	1·55	9·8 ,,
Lead fluoride	1·76	11·1 ,,

Probably the most interesting of the above materials is lead fluoride, since it has a high dispersion similar to that of Double Extra Dense Flint Glass in the visible region. Unfortunately, it is the most difficult to make.

Crystal growing has now entered a phase where many of the classical crystals found in nature and used in optics—such as fluorite—have been grown, and where new crystals not previously found in nature are being successfully made. These may be better for some purposes than the classical crystals. At this stage it is fitting to remember the late Professor Donald C. Stockbarger of the Massachusetts Institute of Technology. He was an outstanding pioneer in the growth of inorganic crystals from the melt, and it is deeply unfortunate that he should have died at an early age just as the field was widening as a result of his efforts. All the above crystals have been grown using his methods, and the help that he freely gave us saved our own work from early frustration.

Department of Natural Philosophy,
University of Aberdeen.
2nd September 1952.

D. A. JONES.
R. V. JONES.
R. W. H. STEVENSON.

Resistance of Germanium Contacts

Previous attempts to explain the high-voltage current–voltage characteristics of germanium point contacts have not taken account of the very high current densities which normally occur near the contact, and which give rise to electric fields far greater than those observed by Ryder and Shockley (1951) to cause non-linear mobility effects of the type discussed by Shockley (1951). The following very approximate theory shows that these effects, together with the Zener current observed by McAfee *et al.* (1951), can account for the observed characteristics, and in particular for the high resistances recently observed for p–p germanium–germanium contacts (Granville *et al.* 1952).

To obtain a simple model for the contact we use an approximation to the observed relation between the field E and current density $i=\pm evn$. The drift velocity v is assumed to be proportional to E up to the field E_1, and constant thereafter, while the carrier density n is assumed to have its normal value up to the field $E_2>E_1$, at which it increases without limit (Zener 1934). The relation between current density, low field conductivity σ_0, and field is thus :

$$i=\sigma_0 E\ (|E|\leqslant E_1); \qquad i=\sigma_0 E_1\ (E_1|<E|<E_2); \qquad i\rightarrow\infty \text{ as } E\rightarrow E_2. \quad \ldots\ldots (1)$$

The geometry is simplified by assuming that the contact between a metal or germanium point and the initially plane surface of a germanium specimen occurs over a hemispherical surface $r=r_0$. For generality it is assumed that there is somewhere a rectifying barrier, but that it has broken down and merely adds a constant voltage V_0 to the voltage drop in the body of the specimen. On applying the relation (1) to the flow field, it is found that for currents I above a certain value I_0, the field is divided into two regions by the hemisphere $r=r_1$. Outside this surface $E=i/\sigma_0$, while within it $E=E_2$. The voltage V at the contact is then

$$V=(E_1/2\pi)^{1/2}(E_2/E_1+1)(I/\sigma_0)^{1/2}-r_0E_2+V_0. \quad \ldots\ldots (2)$$

For $I=I_0$, $r_1=r_0$, and for lower currents the behaviour is everywhere ohmic.

Since the theory is not very sensitive to small departures from the scheme (1), the neglect of the $E^{1/2}$ region found by Ryder and Shockley and the bold extrapolation to constant current throughout the intermediate range may be justified. We take $E_1=1{\cdot}7\times10^3$ v cm^{-1} at 300° K to give the saturation value of v observed by Ryder and Shockley, and $E_2=2\times10^5$ v cm^{-1}.

Some immediate consequences of this relation are :

(i) For voltages rather less than the turnover voltage V_T, a plot of $I^{1/2}$ against V should give a straight line of slope $(E_1/2\pi\sigma_0)^{1/2}(E_2/E_1+1)$. Measurements have been made on the inverse characteristics of five commercial germanium diodes of different types, and on that of the contact between a pointed tungsten wire and the etched surface of an n-type germanium specimen of conductivity $0{\cdot}16\ \Omega^{-1}\,\text{cm}^{-1}$. For three of the diodes, and for the experimental contact under pressures which permit reproducible results to be obtained, a linear relation is found (within the experimental error of 2%) over the range from below 10 v to about $0{\cdot}6V_T$. At higher voltages a more rapid increase occurs which is attributed to the onset of the thermal effects associated with turnover. The slopes vary, in the case of the experimental contact, from 1·1 to 2·0 times that calculated from eqn. (2). Deviations of 10% and 30% are found in the remaining diodes, the larger value occurring in a heavily electroformed type.

(ii) If it is assumed that similar contacts have the same power dissipation at turnover, they should have $V_T\propto\sigma_0^{-1/3}$. Douglas and James (1951) give a curve relating V_T and arsenic content (to which the conductivity should be proportional). Their data, replotted on a log–log basis, do in fact give a nearly straight line of slope $-0{\cdot}35$ for arsenic content greater than $1{\cdot}5\times10^{-6}$.

(iii) Unlike the results of theories involving potential barriers, eqn. (2) shows that, at constant V, $I\propto\sigma_0$, so that a transistor using a contact as collector should have a current gain α not greater than $b+1=3{\cdot}1$. If there is hole trapping, b and α can be greater (Sittner 1952).

(iv) Near room temperature σ_0 and E_2 should be nearly independent of temperature, as should V_0 if it involves the Zener process. Since $E_2\gg E_1$, I should then be proportional to E_1. Shockley (1951) gives low-temperature data on the variation of E_1 which suggest that this effect is of the right magnitude to account for the observed variations in I.

Finally, if it is assumed that the Fermi level at a germanium surface lies near the valence band, a contact on n-type material should be associated with a high barrier and an inversion layer which can cause hole injection. This will cause an increase in σ_0 near the contact for forward current flow, preventing the occurrence of the effects discussed here and ensuring a low resistance. In the inverse direction these effects are adequate to account for the observed voltages even in the absence of any barrier. At a p-type surface, on the other hand, a low barrier without inversion layer will occur and electron injection will be absent. The resistance will then be high in both directions regardless of the nature of the contacting material.

I am indebted to the Directors of Elliott Brothers (London) Ltd., for permission to publish this letter, and to Mr. F. D. Robinson for help with the measurements.

The Research Laboratories of
Elliott Brothers (London) Ltd.,
Elstree Way,
Borehamwood, Herts.
27th August 1952.

J. B. Gunn.

Douglas, R. W., and James, E., 1951, *Proc. Inst. Elect. Engrs.*, **98**, Part 3, 157.
Granville, J. W., Henisch, H. K., and Tipple, P. M., 1952, *Proc. Phys. Soc.* B, **65**, 651.
McAfee, K. C., Ryder, E. J., Shockley, W., and Sparks, M., 1951, *Phys. Rev.*, **83**, 650.
Ryder, E. J., and Shockley, W., 1951, *Phys. Rev.*, **81**, 139.
Shockley, W., 1951, *Bell. Syst. Tech. J.*, **30**, 990.
Sittner, W. R., 1952, *Proc. Inst. Radio Engrs.*, **40**, 448.
Zener, C., 1934, *Proc. Roy. Soc.*, **145**, 523.

The Zernike Circle Polynomials Occurring in Diffraction Theory

In his paper on the phase contrast test Zernike (1934, see also Zernike and Brinkman 1935) introduced certain polynomials $U^m{}_n(x, y)$ which form a complete orthogonal set for the interior of the unit circle $x^2+y^2 \leqslant 1$. Apart from their usefulness in the theory of the phase contrast, these polynomials play a fundamental role in some recent researches concerning the diffraction effects of optical aberrations (Nijboer 1942, 1947, Zernike and Nijboer 1949). Now it can be easily seen that there exists an infinite number of complete orthogonal sets of polynomials for the interior of the unit circle. It is, therefore, natural to enquire about the properties which distinguish the polynomials of Zernike from the other sets. It is the purpose of this note to point out that a uniqueness theorem on Zernike polynomials may be formulated and that with the help of this theorem and from some well-known properties of the Jacobi polynomials explicit expressions for the Zernike polynomials and their corresponding generating function may almost immediately be written down.

It may be shown that a continuous function $V(x, y)$ will be invariant in form with respect to rotations of axes about the origin $x=y=0$ if and only if, when expressed in polar coordinates r, ϕ, it is of the form

$$V(r\cos\phi,\, r\sin\phi)=R(r)e^{il\phi}, \qquad \ldots\ldots (1)$$

where R is a function of r alone and l is an integer (positive, negative or zero). Further it is easily seen that (1) will be a polynomial in x and y of degree n, provided that R is a polynomial in r of degree n, which is such that it contains no powers of r lower than $|l|$ and which is an even or odd polynomial according as l is even or odd. With the help of this result we have established the following *uniqueness theorem:*

There is one and only one* set which is (*a*) orthogonal for the interior of the unit circle, (*b*) contains only such polynomials that are invariant in form with respect to rotations of axes about the origin $x=y=0$, and (*c*) contains a polynomial for each pair of the permissible values of n (degree) and l (angular dependence), i.e. for integral values n and l such that $n \geqslant 0$, $l \gtreqless 0$, $n \geqslant |l|$, and $n-|l|$ is even.*

---

* Polynomials which are such that one is a constant multiple of the other are regarded here as identical.

Let us denote a typical polynomial of this set by V^{l}_{n} and the corresponding radial part by R^{l}_{n}. It may then be shown that $R^{-l}_{n} \equiv R^{l}_{n}$, so that the polynomials may be written in the form

$$V^{\pm m}_{n}(x, y) = R^{m}_{n}(r) e^{\pm i m \varphi}, \qquad \ldots\ldots (2)$$

m denoting non-negative integers. From the properties of the Zernike set and from our uniqueness theorem it immediately follows that the Zernike circle polynomials U^{m}_{n} are simply the real polynomials $R^{m}_{n}(r) \cos m\phi$ and $R^{m}_{n}(r) \sin m\phi$ associated with (2).

Further it follows from the uniqueness theorem that the radial polynomials $R^{m}_{m}(r)$ $R^{m}_{m+2}(r)$, $R^{m}_{m+4}(r) \ldots\ldots R^{m}_{m+2\kappa}(r) \ldots\ldots$ are nothing but the functions obtained by orthogonalizing the sequence r^{m}, r^{m+2}, r^{m+4}, $\ldots\ldots r^{m+2\kappa}$, $\ldots\ldots$ with the weighting factor r over the range $0 \leqslant r \leqslant 1$. Now the well-known Jacobi (or hypergeometric) polynomials may be defined as the functions obtained by orthogonalizing the sequence $1, x, x^2, \ldots\ldots x^{\kappa}, \ldots\ldots$ with the more general weighting factor $x^{q-1}(1-x)^{p-q}$ $(q>0,\ p-q>-1)$ over the range $0 \leqslant x \leqslant 1$. Using this result, it is easily seen that with suitable choice of p and q and with the help of a simple algebraic transformation, the explicit expressions for the R polynomials may immediately be obtained from those for the Jacobi polynomials. Similarly one can obtain the corresponding generation function. If we choose the same normalization as Zernike, namely $R^{m}_{n}(1)=1$ for all m and n, the generating function is found to be

$$\frac{[1+z-\{1+2z(1-2r^2)+z^2\}^{1/2}]^{m}}{(2zr)^{m}\{1+2z(1-2r^2)+z^2\}^{1/2}} = \sum_{\kappa=0}^{\infty} z^{\kappa} R^{m}_{m+2\kappa}(r). \qquad \ldots\ldots (3)$$

A detailed discussion concerning the subject-matter of this note will be published elsewhere.

We are indebted to Professor Max Born and Dr. H. Schneider for helpful discussions.

Department of Mathematical Physics,
University of Edinburgh.
24th September 1952.

A. B. Bhatia.
E. Wolf.

Nijboer, B. R. A., 1942, *Thesis*, University of Groningen; 1947, *Physica*, **13**, 605.
Zernike, F., 1934, *Physica*, **1**, 689.
Zernike, F., and Brinkman, H. C., 1935, *Proc. K. Acad. wet. Amst.*, **38**, No. 2.
Zernike, F., and Nijboer, B. R. A., 1949, Contribution in *La Théorie des Images Optiques* (Paris : Éditions de la Revue d'Optique), p. 227.

---

* This theorem is analogous to the uniqueness theorem on Legendre polynomials, i.e. the set of Legendre polynomials $P_n(x)$ is the only set which is orthogonal for the interval $-1 \leqslant x \leqslant 1$, and which contains a polynomial of every degree.

Optical Constants by Reflection

A recent paper by Avery (1952) briefly reviews methods used to determine the optical constants n and k of a substance by reflection and describes a new method based on measurements of reflection coefficient ratios at various angles of incidence. This, or any other method essentially involving oblique incidence, may be difficult to apply to anisotropic materials, especially those which are biaxial. The following theory has therefore been developed in order to calculate values of the complex refractive index $\mathbf{n}=n-ik$ from reflection coefficient measurements taken at normal incidence over a range of frequencies. It has been applied in the first instance to isotropic materials but can easily be extended to cover crystals of orthorhombic and higher symmetry.

The theory is based on the linearity of the phenomena concerned. Such ideas have been developed extensively in connection with electric network theory (Bode 1945), and it has been shown, *inter alia*, that the attenuation and phase at the output of an electric network are not independent of one another. If one of these is known as a function of frequency the other can be calculated at any frequency. In the absence of ferroelectric and ferromagnetic materials Maxwell's electromagnetic equations are linear. It is then possible to deduce the phase shift upon reflection of an electromagnetic wave at the surface of a semi-infinite slab of material from a knowledge of the modulus of the complex reflection coefficient $\mathbf{r}$ as a function of frequency. That is to say measurements of reflecting power $|\mathbf{r}|^2$ or reflection attenuation $-\ln|\mathbf{r}|$ over a reasonably wide range of frequency can be used to calculate the phase shift at any frequency within the range. The complex reflection coefficient $\mathbf{r}$ is then known in modulus and phase. From it the complex refractive index $\mathbf{n}$ may be completely determined through the reflection equation for normally incident light $\mathbf{r}=(\mathbf{n}-1)/(\mathbf{n}+1)$.

As an example of this method the optical constants of unoriented polythene have been determined in the range 2600–3100 cm^{-1}. The apparatus was designed to deal with small samples of material and it incorporated a reflecting microscope of N.A.$=0\cdot3$. The sample of polythene, approximately a cube of 3 mm edge, was placed at the focus of the microscope. Measurements of reflecting power $|\mathbf{r}|^2$ were made using a Grubb Parsons infra-red spectrometer fitted with a large LiF prism and a Hilger-Schwarz thermocouple which was connected to a General Motors d.c. amplifier. The monochromator half-energy slit width was 7 cm^{-1}. The graph of reflection attenuation against frequency was split into 17 finite straight line segments after the manner suggested by Bode (1945, p. 337) in connection with the attenuation of an electric network, and charts similar to his were used to compute the phase contributions of each line segment. It is interesting to note that a constant percentage error in the measurement of $|\mathbf{r}|^2$ does not influence the phase calculations. As the N.A. of the microscope set a lower limit of about 22° for the angle of incidence of an axial ray, a correction for obliquity was applied so that $\mathbf{r}$ was known for normally incident light. These $\mathbf{r}$ values were plotted on an Argand diagram (fig. 1). The reflection equation $\mathbf{r}=(\mathbf{n}-1)/(\mathbf{n}+1)$ shows that the loci of points of constant n and constant k consist of two families of circles. A circle for n constant has radius $1/(n+1)$ and is tangential to the ordinate at $\mathbf{r}=(1, 0)$ and a circle for constant k has radius $1/k$ and is tangential to the abscissa at $\mathbf{r}=(1, 0)$. From this diagram the refractive index n and the extinction coefficient k can be read off directly. The values so obtained are plotted against frequency in fig. 2. An absorption measurement with a film of the same polythene 2·0 microns thick gave peak values of $k=0\cdot25$ and $0\cdot42$ at 2850 and 2920 cm^{-1} respectively, in good agreement with $k=0\cdot24$ and $0\cdot41$ at 2848 and 2922 cm^{-1} respectively, deduced from the reflection spectrum.

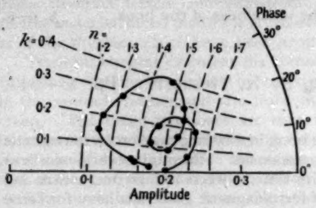


Fig. 1. Polar plot of complex reflection coefficient for polythene in the range 2600–3100 cm^{-1}. Curves of constant n and k are superimposed and shown as broken lines.

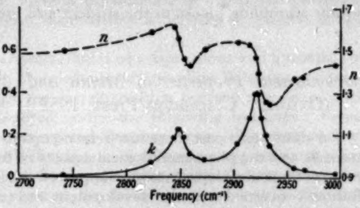


Fig. 2. Refractive index n and extinction coefficient k for polythene from 2700 to 3000 cm^{-1} as calculated from reflection spectrum.

Thanks are due to Professor J. T. Randall for research facilities, to Dr. W. C. Price for his interest in this work, to the Department of Scientific and Industrial Research for a maintenance grant and to the Institute of Petroleum for financial assistance.

Wheatstone Physics Laboratory,
King's College, London.
19th September 1952.

T. S. Robinson.

Avery, D. G., 1952, *Proc. Phys. Soc.* B, **65**, 1087.
Bode, H. W., 1945, *Network Analysis and Feedback Amplifier Design* (New York : Van Nostrand).

REVIEWS OF BOOKS

Sound Insulation and Room Acoustics, by PER V. BRÜEL, (translated by J. M. BORUP). Pp. xi + 275. (London: Chapman and Hall, 1951). 35*s*.

Quite a number of books on sound insulation and room acoustics have appeared in recent years, and the reviewer naturally looks first for new information and originality of treatment. On the first point this book presents information obtained in the last few years in Denmark and Sweden, where the author's researches have been made, and not previously available in English. Though this work to some extent overlaps work done in Britain, Germany, Holland and U.S.A., new and interesting data are presented, especially on the properties of resonant absorbents and compound walls, while Dr. Brüel's considerable experience in the acoustic design of Scandinavian buildings is placed at the disposal of the reader. The nomograms for design data should be very useful to the worker in this field.

One cannot however praise the presentation, and in this one feels that Dr. Brüel is not well served by his translator. No fault indeed can be found with the grammar, but the translator does not appear to be well acquainted with English terminology in acoustics, and often plays for safety by a too literal rendering. Thus, we do not in this country speak of ' angular frequency ', ' completely tangent sound wave ', and ' different in-swinging conditions', while in England musical intervals are measured in cents, not in centimetres (cf. Table II). Such peculiarities, while they are fairly easily resolved by the expert, may well become stumbling blocks to the student and limit the usefulness of the book. E. G. R.

Ferromagnetic Properties of Metals and Alloys, by K. HOSELITZ. Pp. xi + 317. (Oxford: Clarendon Press, 1952.) 40*s*.

For some years past there has been no up-to-date book in any language on ferromagnetic materials and the phenomenological aspects of ferromagnetism. Although several excellent reviews of new developments have appeared during this time, workers in ferromagnetism and all those concerned with the development and use of ferromagnetic materials have for some time felt the lack of a comprehensive account. Since, in addition, the last decade has been a period of great activity in all branches of the subject, it is scarcely surprising that now, not only in this country but in the rest of Europe and in America too, several books on ferromagnetism are appearing. To some extent all deal with the same subject matter, but the inevitable differences of approach mean that there is not a great deal of overlapping and also that it is of some interest to consider the points in which they differ from each other.

With this in mind, perhaps the most outstanding feature of *Ferromagnetic Properties of Metals and Alloys* is conveyed immediately by the title since, quite deliberately, the book deals only with metals and alloys and does not discuss at all one of the most spectacular post-war advances—antiferromagnetism and non-metallic magnetic materials. Until recently Director of Research to the Permanent Magnet Association, Dr. Hoselitz is well aware that many of the problems of magnetic alloys are of metallurgical origin, and his book is written with special emphasis on their physical metallurgy. It includes a chapter on the general properties of metals and alloys as a background to the main subject matter, and also a chapter on magnetic analysis. This is probably the first complete account of the ways in which a study of magnetic properties, particularly (σ, T) curves, may be used to provide information on the constitution of ferromagnetic alloy systems.

The author has achieved a nice balance between theory and practice, although the properties of soft magnetic materials in particular are discussed from a rather academic point of view, with relation to their metallurgical constitution rather than to their technological applications. As might be expected, there is, of course, a wealth of data on hard magnetic materials. The theoretical background is explained adequately but not in great detail, and again it is a little surprising to so find little mention of the recent work on domain shapes and sizes.

There are very few errors in the text, but it should be pointed out that the magnetic sorting bridge does not primarily measure hysteresis loss and that the loop observed on a cathode-ray oscillograph from a specimen excited at 50 c/s is not a hysteresis loop as ordinarily understood. With the exception of the captions to figs. 47 and 72 and a slight slip at the bottom of p. 179 no misprints have been detected, and there is an adequate and carefully compiled index. A. E. DE BARR.

Creep of Metals, by L. ROTHERHAM. Pp. 80. (London: Institute of Physics, 1951.) 15*s*.

The present-day demands for materials for use at high temperatures have served to focus attention on the manner in which the mechanical properties of metals depend upon temperature. In particular, the creep deformation, which in most metals is insignificant at normal temperatures but is of great importance at elevated temperatures, has been studied extensively. Nevertheless, while our knowledge and understanding of deformation processes in metals have advanced rapidly during the last quarter of a century, the problem of creep still presents a somewhat confused picture in that there exists, as yet, no generally accepted unified theoretical treatment. A number of different processes may contribute to creep deformation, so it is not surprising that attempts to formulate such a theory applicable to all metals have been unsuccessful. There has accumulated a mass of experimental information somewhat difficult to correlate and a number of theoretical approaches have been suggested. To bring to readers a clear and concise account of creep as our knowledge of it now stands is not an easy task and the author's attempt to do so within the narrow compass of eighty pages deserves much commendation. The general phenomena of creep are presented clearly and a balanced picture is given of the basic mechanisms underlying the process as well as of the theories which have been proposed.

The first chapter discusses the common characteristics of creep curves and a number of empirical relations which have been derived. The elements of crystalline flow are considered in the second chapter and the third outlines the metallographic features associated with creep. The creep processes are then considered under the following headings: Grain boundary creep, Transient creep, Steady state creep and Tertiary creep. The final chapter gives a brief though valuable indication of the difficulties in developing creep resistant alloys and of some of the directions in which improved creep resistant materials may be sought.

The book is written from the point of view of the metal physicist with a thorough appreciation of practical metallurgists' problems, and should serve as a valuable introduction for both physicist and metallurgist to the significant publications in this field, most of which the author has mentioned. RONALD KING.

Les machines à penser, by LOUIS COUFFIGNAL. Pp. 153. (Paris: Les Editions de Minuit, 1952.) No price given.

The author defines a thinking machine as a machine which can replace a human being in the performance of operations belonging to some class, however limited. Calculating machines clearly come into this category and their development is described beginning with the work of Pascal and ending with modern large-scale digital and analogue machines. Some comparisons are drawn between calculating machines and the cerebral cortex, both in relation to action and structure. This leads to a discussion of symbolic logic and ways of mechanizing it, and finally to a chapter on the problem of thought—but the net result is, as the author admits in a postscript, the statement of a series of questions to be answered by future research rather than a conclusion.

The author's approach to the problem of thought is to consider the brain as a machine which can derive systematically the consequences of propositions expressed in symbolic form and held in its memory, and which can store these consequences and use them as new data for a further series of deductions. He considers that the logical rules followed may not be the same as those of Aristotelian logic, and that light may be shed on them by a statistical examination of texts in a wide range of subjects in order to see what concepts are normally linked together. A diagram is given showing how certain logical processes can be carried out in a machine composed of elementary logical and storage elements. A comparison—

which may perhaps strike some readers as a little naïve—is made between the connections in this machine and the arrangement of nerve fibres in the brain. The philosophical aspects of the subject are not embarked on.

The author appears to have been less influenced by a consideration of the properties of large universal calculating machines than have other writers who have stressed the ability of such machines to discriminate between different kinds of stimuli received through their input mechanisms and to modify their own programmes of instructions accordingly. In fact the amount of space devoted to universal calculating machines is sufficient for little more than a general description.

It would have been helpful if more references had been given to the sources on which the arguments are based, and if an index had been provided. Apart from this, however, the book is well written.

M. V. WILKES.

Applied Statistics Vol. 1, No. 1. (A new journal of the Royal Statistical Society.) Edited by LEONARD H. C. TIPPETT. Pp. 80. (London: Oliver and Boyd, 1952.) 10*s*. Annual subscription 25*s*. per annum.

One of the sad facts of life is that pressing needs often provoke depressing remedies; but in this first number of *Applied Statistics* Mr. Tippett and his colleagues have unerringly diagnosed the need and expertly compounded the remedy. Their triumph is presented in elegant format by Messrs Oliver and Boyd.

The social value of scientific knowledge fructifies in applied science where the importance of careful scientific methodology is at least as important —and often more difficult—than in the laboratory of the fundamental scientist. Pure science thrives in the controlled conditions of the laboratory where standardization of interfering variables reduces experimental error and thereby renders more easy the detection of significant relationships in the experimental data. In the applied and social sciences, however, considerations of economy and fundamental human freedoms frequently preclude any close approach to standardized conditions. Bias has to be shattered by the device of randomization. The importance of variables which will interfere in the field has to be estimated, with a consequent spreading of experimental error which entails relatively extensive experiments and the use of powerful statistical tests of significance. The mechanics of experimentation become so complex that it is not surprising that recent decades have seen experimental design grow into a formal discipline with its own principles and technique that demand very careful study.

Mathematical statistics and applied statistics have continuously prodded each other into an unprecedented torrent of usefulness, but the inherent difficulty of many of the problems and the esoteric formulation of their solutions demands an interpreter, since it is inevitable that the greatest practical benefits of statistics arise when the techniques are applied by non-statisticians to their own experimental designs and analyses. Statistics is the science of the collection and analysis of data. It is invaluable, therefore, in all branches of science. For this reason we say without hesitation that *Applied Statistics* should be regularly available to every scientific worker, a discouraging number of whom are sadly in need of the assistance it sets out to offer. The fact that the physicist will see in it articles dealing with the techniques used in chemistry, physiology, market research, coal mining, accountancy, psychology and so on is all to the good. It will be invaluable to him to stand back from physics and consider the analysis of data as such. When, later, he finds an article on " The Calibration of a Gas Flowmeter " he will appreciate it all the better.

Industrial physicists who have an Industrial Applications Group of the Royal Statistical Society in their vicinity would do well to join; not only for the professional benefits to be derived, but also because members of such groups can deduct 10*s*. from the subscription rate to this journal, and the fee for membership of a group is only 10*s*. 6*d*. per annum. The only qualification for membership of a group is interest.

There is no doubt that this journal will very soon be one of the most useful and popular available to us. Mr. Tippett and his colleagues have done us great service by their enterprise. Demand for the journal will soon convince them of our gratitude.

M. J. MORONEY,

La Cybernétique : Théorie du Signal et de l'Information (Réunions d'Études et de Mises au Point tenues sous la présidence de Louis de Broglie). Pp. vi+318. (Paris : Éditions de la Revue d'Optique, 1951.) 1600 fr.

This is the collection of a series of lectures, by 14 invited lecturers, delivered in Paris in the spring of 1950 at the Ecole Polytechnique under the chairmanship of Louis de Broglie. Their purpose was to give French scientists and engineers a comprehensive picture of the new science which in this country and in the United States is known as Communication Theory. It is somewhat unfortunate that M. Julien Loeb, the organizer, has chosen the name " Cybernetics ", a term which means " the art of aiming or steering ", and which was first used by Norbert Wiener for what engineers used to call by the humbler name of " theory of closed loop systems ". There is in fact not a single lecture in this series concerning ' cybernetics ' in this original meaning of the term, but the damage is done. In France Communication Theory has been henceforth known as ' Cybernétique '. This however is the only damage, hardly worth mentioning in view of the good done by this lecture series, which was soon followed by a remarkable crop of original contributions, putting French workers into the front rank especially on the more theoretical side of the subject.

At the time of the lecture series, only two years ago, French workers had made comparatively few original contributions and the chief purpose was surveying the work done mainly in the United States and partly also in this country and in Russia. The reviewer was the only foreign contributor. The French authors have so well succeeded in their task that even now, after two years, this book can be recommended as the best comprehensive survey of this field. (The British Symposium on Information Theory, in September 1950, organized by Professor W. Jackson, covered an even wider domain, but it aimed at original contributions rather than at completeness.)

In the first lectures R. Fortet reviews Markoff chains, Indjoudjian Wiener's filtering theory, Black–Lapierre noise theory, and Aigrain Shannon's work. These are followed by Oswald's lecture on analytical signals, Gabor's on communication theory and physics and Ville's interesting theoretical discussion of the information concept, which he has greatly extended in the meantime in collaboration with Schutzenberger. The later lectures are of a more practical and engineering nature. Chavasse spoke on telephony, Colombo on distortion, Delbord on television, Icole on multiplex, Marcou on frequency compression, and Picault on coding.

It may be hoped that this volume, which has so much benefited French science and engineering, will obtain in this country as well the attention which it so rightly deserves.

D. GABOR.

Polarized Light in Metallography, by G. K. T. Conn and F. J. Bradshaw. Pp. xi+130. (London: Butterworths Scientific Publications, 1952.) 21*s*.

This monograph was prepared under the auspices of the Optical Methods Sub-Committee of the British Iron and Steel Research Association in order to survey the field of application of the reflecting polarizing microscope to the problems of the metallurgist. Much more interest is now being shown in this technique, and many development have been made in recent years. This book gives an authoritative assessment of the present position and of future prospects, a number of well known workers in this field being responsible for the different topics.

The first two chapters introduce the basic ideas of the relevant parts of physical optics in sufficient detail for a beginner to follow the remaining chapters of the book, although a fuller and deeper understanding and appreciation of the subject requires more extensive reading of the books given in the bibliography. Quite correctly in a book of this kind results are quoted without proof and emphasis is laid on the meaning and application of the formulae, illustrated by graphs and experimental results. It is a pity that a number of errors and ambiguities should have crept into what is otherwise an admirable introduction. For example mica and selenite on pp. 12 and 14 are incorrectly termed uniaxial crystals. " Opaque materials such as metals " on p. 5 should surely read " Transparent materials such as glass ". Unnecessary repetition in the second paragraph on the quarter wave plate is liable to confuse the reader. " Quartz " should clearly be omitted from the first sentence of the middle paragraph on p. 16, and the " or " in (bound or free electrons) should of course be " and ".

Chapter 3, on ' Equipment and procedure for microscopy by reflected polarized light ', is very comprehensive and complete and obviously written by masters of their subject. Together with the references given it should prove a great help to those equipping themselves for this work. Great care is taken to explain the procedure necessary for exact adjustment of the microscope and for the methods of use so as to obtain unambiguous results.

The next chapter is concerned with the detailed application of the technique to the examination of metal surfaces. The problems very ably dealt with are : (*a*) Revealing grain structure, (*b*) Detection of preferred orientation, (*c*) Examination of oxide coatings, (*d*) Detection of internal strains and plastic deformation, (*e*) Identification of phases in multiple structures. This list shows what a powerful tool the polarizing microscope can be when used intelligently. There are of course many difficulties, and care is necessary to avoid erroneous conclusions. These are explained in detail and many examples are given, supported by over 60 references. There is one printing error. " Wilkensen " on pp. 81 and 89 should be " Wilkinson ".

Chapters 5 and 6, on the examination and identification of inclusions, and on the study of ores, are very good and complete, with much useful information and with comprehensive references. Particular care is taken to guide the beginner away from pitfalls and false conclusions.

The book ends with a short but useful summing up and a glossary of optical terms. For metallurgists interested in using the reflecting microscope this book will be found invaluable. J. BOR.

X-Ray Crystallographic Technology, by A. GUINIER. Pp. xiii+330. (London : Hilger and Watts, 1952.) 56*s*.

The eloquent appeal issued by Professor Lonsdale when reviewing the original French edition (1945) in this journal has now been answered by this production of a revised " English text in a better setting ".

The few points which aroused unfavourable comment have all received attention (although the cover has altered only in shade and not in colour !). Probably the most important and useful alteration is the addition of two indexes, a subject index and a combined name and bibliography index. The English title, in contrast to *Radiocristallographie*, gives a clear indication that the book is primarily intended as a working manual for those using x-ray diffraction as one among several methods employed on a research problem. The amount of basic theory included, and the manner in which it is presented, in fact make it a useful textbook for all students of x-ray crystallography. Particularly helpful are the introductory and summarizing statements before and after each section, and the way in which many points usually ignored by textbooks, but puzzling to the uninitiated, are dealt with.

The scope of the book, very wisely, does not include methods of determining crystal structures, but some examples of known structures are given, and it is shown how the intensities of the x-ray reflections from such structures can be calculated. Unfortunately, the treatment of these matters is such that two long-standing confusions have once more been given a new lease of life. The first, which occurred also in the original edition, concerns the distinction between the Bravais lattice and the atomic structure. This distinction is not clearly made and the position is not helped by the use of terms such as ' the structure of the point lattice ', and the application of the term ' simple lattice ' to a structure containing one atom per lattice point. The second confusion has been introduced by the indiscriminate use of the terms ' structure factor ' and ' structure amplitude ' for the same expression.

Some other slightly misleading words or phrases, mostly more amusing than serious, arise from a rather too literal translation. For instance the term ' crystal with base ' does not in fact mean a museum exhibit provided with a glossy stand.

The practical section, which constitutes more than half the book, is concerned with experimental methods and their applications, including difficulties and limitations. By an admirable choice of relevant data, examples, and actual photographs, it succeeds in making the practical aspects as 'real' as any textbook can hope to do. The review of experimental methods is not intended to be exhaustive, and it is only to be expected that every x-ray crystallographer will find that some of his favourite practical hints are not included; one such noticeable omission is the use of goniometer slides for centring a single crystal.

The defects cited above are such that they do not seriously detract from the merits of the book, and it can be confidently recommended as a useful addition to the libraries of all interested in the subject. The 'better setting' appealed for has certainly been realized, and it only remains to lament that so high a price has to be paid for a book which is of so much potential use to so many students. AUDREY M. B. DOUGLAS.

Color in Business, Science and Industry, by DEANE B. JUDD. Pp. ix+401. (New York: Wiley; London: Chapman and Hall, 1952.) $6.50, 52*s*.

The number of books on the scientific and technical aspects of colour is steadily increasing, so that the discriminating reader is now in the happy position of being able to choose which volumes he shall have on his library bookshelf. This latest addition must certainly be included in the list, for Dr. Judd is an outstanding authority on colour and he has written an extremely informative and stimulating book.

The first things one notices are the enthusiasm of the author for his subject and the pride which he takes in American contributions to it. Twice in the course of a short introduction he assures us that "American colorimetry is the envy of the rest of the world" and while this country played an active part in the establishment of an international system of colour measurement, it is only too true that American exploitation of the system has been far more extensive than in Great Britain.

Another feature that soon becomes apparent is that this is a practical book, not necessarily in the sense that detailed descriptions of how to prepare samples or how to operate a piece of equipment are given, but because the commercial value of colorimetric data is constantly in the mind of the author and is frequently being pointed out to the reader. This value has been brought home to Dr. Judd through his experience in helping to solve many industrial problems about which his advice has been sought.

The book is divided into three parts. Part I deals with the basic facts of colour vision, colour mixture and matching and with the different approaches to colour, from that of the physicist to that of the customer; Part II is by far the largest section of the book and covers the C.I.E. system of colorimetry, the design of spectrophotometers and colorimeters, differential colorimetry, the colour solid and colour atlases, colour reproduction, and so on; while Part III discusses colorant layers, including the application of the Kubelka–Monk analysis to the calculation of the reflection characteristics of dyed and pigmented surfaces.

The absence of separate chapters may prove a little confusing to some readers, although there is usually a natural transition from one section to the next. Occasionally the arrangement of the subject-matter appears a little illogical, as when a section on the Reproduction of Pictures in Color is sandwiched between Visual Colorimeters and Colorimetry by Difference. On the other hand, the various colour atlases are related to one another and to the colour solid in a most instructive manner and anyone wishing to know the strong and weak points of each system could hardly do better than consult Dr. Judd. To many people Part III will be valued the most highly because the subject of colorant layers is of great technical importance and has not been adequately dealt with in other books on colour.

Many points could be picked out for specially favourable notice if space permitted; for example, the discussion on p. 119 of the limits of accuracy of the selected ordinate method of calculating tristimulus values; the remark on p. 36 that there is no need to involve fancy geometry in discussions of colour space; the homely story of Jack and Jill on p. 278 which provides a useful warning to the too-clever psycho-physicist; the recognition on p. 270 that, while colour tolerances can be expressed scientifically, the limits that can be tolerated will depend on whether you are operating in a buyer's or a seller's market; and many others besides.

Dr. Judd has always been noted for his interest in work outside the United States, and this book is no exception. It is pleasing, for example, to find the Donaldson 6-stimulus colorimeter given adequate description; there is a lengthy section on the Tintometer, with an interesting historical account of how the Tintometer glasses had such a widespread use in America; the Terminology Report of the Physical Society Colour Group is referred to and Duncan's work on pigment mixtures earns several references; there is a very fair account of the Villalobos Colour Atlas produced in the Argentine; Pitt's chromaticity confusion loci for dichromats are shown, although no reference is made to Pitt himself. Presumably the section on spectrophotometers was written at too early a date for any reference to be made to the Cary Recording Spectrophotometer developed in California and now used in a number of laboratories in America.

The book appears to be free from any obvious errors, although on p. 73 the statement that " there are probably less than half a million totally colour-blind people in this country (U.S.A.) ", while not untrue, hardly gives the right impression. It looks as if there are a couple of noughts too many. A rather unhappy diagram of a spectrometer appears on p. 81, an unfortunate inheritance from Evans' *Color*. Otherwise the book is well, if not profusely, illustrated, and the quality of production is up to the high standard we have come to expect in American publications.

W. D. WRIGHT.

CONTENTS FOR SECTION A

PAGE
Prof. Louis Néel. Antiferromagnetism and Ferrimagnetism. (7th Holweck Lecture) . . . 869
Mr. A. B. Lidiard. Spin Degeneracy and the Theory of Collective Electron Ferromagnetism . . . 885
Dr. A. H. Cooke and Dr. L. E. Drain. Proton Magnetic Resonance and Molecular Motion in Solids . . . 894
Dr. R. G. Chambers. The Two-Band Effect in Conduction . . . 903
Dr. J. H. Fremlin and Miss M. C. Walters. An Experimental Investigation of the Stability of Nuclei against Double Beta-Disintegration . . . 911
Mr. H. C. Newns. The Theory of (d, t) Reactions . . . 916
Prof. D. R. Bates and Dr. A. Dalgarno. Electron Capture—I: Resonance Capture from Hydrogen Atoms by Fast Protons . . . 919
Dr. J. R. Prescott. The Transition Effect for Cosmic-Ray Bursts at Small Thicknesses of Lead . . . 925
Mr. L. S. Kothari. On a Modified Definition of Riesz Potential for the Meson Case . . . 930
Prof. C. A. Coulson. Alkyl Shifts in Absorption Spectra of Azulene and other Aromatic Molecules . . . 933
Dr. F. D. S. Butement and Mr. R. Shillito. Radioactive Gold Isotopes . . . 945
Letters to the Editor:
  Dr. R. J. Blin-Stoyle. Spin–Orbit Coupling in the D–D Reaction . . . 949
  Dr. J. H. E. Griffiths and Mr. J. Owen. The Nuclear Spins of ^{99}Ru and ^{101}Ru . . . 951
  Mr. J. M. Baker and Dr. B. Bleaney. Nuclear Spin of Vanadium 50 . . . 952
  Dr. B. Bleaney and Dr. D. J. E. Ingram. Paramagnetic Resonance in Potassium Ferricyanide . . . 953
  Dr. H. E. D. Scovil and Dr. K. W. H. Stevens. The ' Extra ' Levels in Rare Earth Salts . . . 954
  Dr. A. B. Pippard and Dr. R. G. Chambers. The Mean Free Path of Conduction Electrons in Bismuth . . . 955
  Mr. A. Teviotdale. Zener's Treatment of Ferromagnetism. . . 957
  Dr. G. E. Harrison and Mr. E. D. Seymour. The Thermal Neutron Capture Cross Section of ^{84}Sr . . . 958
  Miss M. C. Walters. The Origins of Background Tracks in Shielded Electron-Sensitive Emulsions . . . 959
Reviews of Books . . . 960
Corrigendum (M. H. L. Pryce) . . . 962
Contents for Section B . . . 963
Abstracts for Section B . . . 963

ABSTRACTS FOR SECTION A

Antiferromagnetism and Ferrimagnetism, by Louis Néel.

ABSTRACT. The present position of our knowledge of antiferromagnetism including ferrimagnetism is reviewed, and some very interesting phenomena concerning the magnetic behaviour of certain ferrites and of pyrrhotite are described and explained.

Spin Degeneracy and the Theory of Collective Electron Ferromagnetism, by A. B. Lidiard.

ABSTRACT. The theory of ferromagnetism is often developed on the assumption that the exchange energy responsible for the magnetization is adequately represented as a sum of exchange integrals between pairs of electrons with mutually parallel spins. This is equivalent to using for the total wave function a determinant built up of one-electron functions. For the magnetized states of interest in ferromagnetism a single determinant wave function must be regarded as inadequate on account of the so-called spin or exchange degeneracy, which requires us to use instead appropriate linear combinations of the degenerate determinants. In the present paper a model is presented for which the results of neglecting or taking account of this spin degeneracy are identical. The model can be handled exactly and leads to equations for the free energy, magnetization, etc., which are a generalization of Stoner's theory of ferromagnetism to a system with two energy bands; these could be the overlapping 3d and 4s bands in the iron transition series. The calculations are not necessarily restricted to ferromagnetic metals and would apply equally well to those that are paramagnetic.

Proton Magnetic Resonance and Molecular Motion in Solids, by A. H. Cooke and L. E. Drain.

ABSTRACT. Proton magnetic resonance widths and relaxation times have been measured in ammonium chloride, bromide and iodide, methyl alcohol and dimethyl sulphate from 90° K to 290° K. Line width transitions were observed in NH_4Cl and NH_4Br at 135° K and 99° K respectively. There were no sharp changes in either the relaxation time or line width at the temperatures of the λ-specific heat anomalies of the ammonium salts. Evidently, there is no marked increase in the rate of rotation of the ammonium ion at these transitions. The results for NH_4Br suggest that rotation of the NH_4^+ ion is hindered by a potential barrier of about 3 500 cal/mole. The N–H distance is $1{\cdot}03 \pm 0{\cdot}01$ Å. The line width and relaxation time show pronounced changes at the λ-point of methyl alcohol, indicating a sudden change in the motion of the CH_3OH molecule. In solid dimethyl sulphate, a narrow line of small amplitude appears at about 170° K.

The Two-Band Effect in Conduction, by R. G. Chambers.

ABSTRACT. For an isotropic two-band conductor in a transverse magnetic field, the various galvanomagnetic, thermomagnetic and thermoelectric coefficients are derived by elementary arguments in terms of those of the individual bands. The expressions are applicable to quasi-free electrons of any degree of degeneracy: for fully degenerate electrons they reduce to those already given by Sondheimer and Wilson. When the square of the thermoelectric power is comparable with the Lorenz number, as may happen in semiconductors, the expressions for the adiabatic thermomagnetic and thermoelectric effects become extremely complicated, though those for the isothermal effects remain reasonably simple.

An Experimental Investigation of the Stability of Nuclei against Double Beta-Disintegration, by J. H. Fremlin and M. C. Walters.

ABSTRACT. Samples containing elements of even Z have been examined for possible beta-activity using a photographic emulsion technique. For the elements Cr, Fe, Ni, Zn, Ge and Cd the results show that the effective half-value period for double beta-decay for the elements as a whole is in each case greater than 10^{17} years if the disintegration energy is 2 MeV or more (or greater than 10^{16} years if the disintegration energy is $0{\cdot}35$ MeV or more). For Ca, Mo, Sr, Sn, Te, Ba, W, Os and Pt lower limits between 2×10^{14} and 6×10^{16} years are similarly obtained on the assumption of 2 MeV disintegration energy, the residual uncertainty in most cases being due to the slight contamination of the samples with natural radio-elements. A possible weak beta-activity shown by samples containing Mo requires further examination.

The Theory of (d, t) Reactions, by H. C. Newns.

ABSTRACT. Expressions for the angular distribution of tritons emitted from nuclei bombarded with deuterons have been obtained using Born's approximation and the non-perturbation method of Butler. Discussion of the effect of different forms of the triton wave function on the angular distribution is given. The theory is shown to explain the experimental angular distributions satisfactorily.

Electron Capture: I—Resonance Capture from Hydrogen Atoms by Fast Protons, by D. R. Bates and A. Dalgarno.

ABSTRACT. Born's approximation is used to calculate the cross section associated with resonance charge exchange between protons and hydrogen atoms, it being pointed out that in earlier treatments of this problem unjustified simplifying assumptions were made. An energy range of up to 250 kev is covered. The predicted energy variation of the cross section is in harmony with the results of Ribe's recent laboratory work on collisions of protons with hydrogen molecules.

The Transition Effect for Cosmic-Ray Bursts at Small Thicknesses of Lead, by J. R. Prescott.

ABSTRACT. The transition curve for cosmic-ray bursts under lead at sea level is analysed to find the contributions from stars, extensive showers, electromagnetic interactions of μ-mesons, and single high-energy electrons and photons. When this is done, a substantial fraction of the bursts under lead thicknesses less than about 2·5 cm remains unaccounted for. It is suggested that they are produced by ' narrow air showers ' consisting predominantly of photons of about $2{\cdot}5 \times 10^8$ ev and having a sideways spread of less than 0·3 m.

On a Modified Definition of Riesz Potential for the Meson Case, by L. S. Kothari.

ABSTRACT. A modified definition of the Riesz potential for the meson field is introduced. It is shown that the new definition is a generalization, in the α-plane, of the meson potential in the interaction representation. The relations satisfied by the meson potential in the interaction and in the Schrödinger representations are compared.

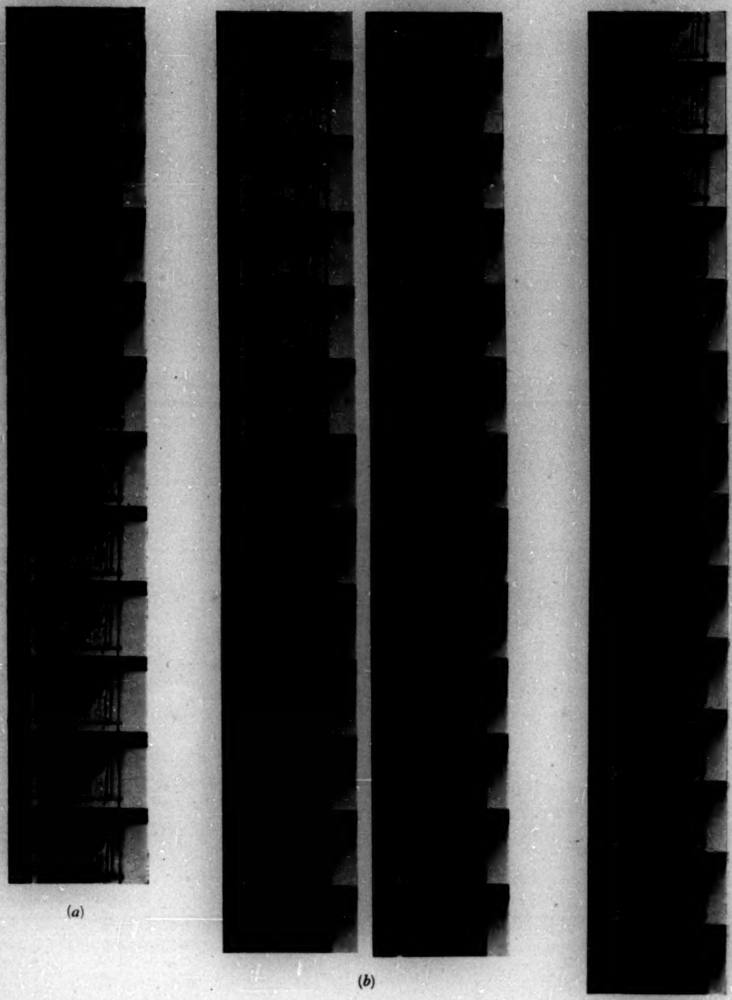
Alkyl Shifts in Absorption Spectra of Azulene and other Aromatic Molecules, by C. A. Coulson.

ABSTRACT. An inquiry is made into the factors which cause shifts in the ultra-violet spectra of aromatic molecules on substitution. The method of molecular orbitals is used, and the discussion is confined to π-electrons. It is shown that there are two chief effects : the first is essentially electrostatic, or inductive ; the second is essentially conjugative, or mesomeric. In alternant hydrocarbons the second predominates and causes a red shift. In other molecules (e.g. azulene, quinoline) the first effect is considerably larger and now predominates over the second. Different positions of substitution will lead to shifts, sometimes to the red and sometimes to the blue. Electron donors will cause shifts in opposite directions to electron acceptors. The importance of these variations in characterizing the electronic transitions is stressed. Detailed investigation of the various methyl azulenes confirms the assignment of the long wavelength band with almost complete certainty.

Radioactive Gold Isotopes, by F. D. S. Butement and R. Shillito.

ABSTRACT. The half-lives, mass assignments and radiation characteristics of three neutron-excess gold isotopes have been determined as follows :

	^{200}Au	^{201}Au	^{203}Au
Half-life	48 min	26 min	55 sec
Max. β-energy (MEV)	2·2	1·5	1·9
γ-energy (MEV)	0·39, 1·13	0·55	0·69



(*a*)

(*b*)

(*c*)

Fig. 4.

.B
5
2